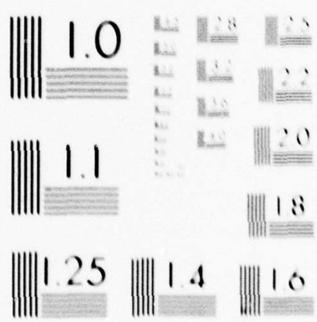
AD-A068 666 YALE UNIV NEW HAVEN CONN DEPT OF COMPUTER SCIENCE F/G 5/7
KNOWLEDGE-BASED PARSING.(U)
APR 79 A V GERSHMAN N00014-75-C-1111
UNCLASSIFIED RR-156 NL

2 OF 2
AD
A068666

END
DATE
FILMED
6-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

attitudes and give rise to their goals are called <u>themes</u> (Schank and Abelson 1977, Wilensky 1978). For example, if we know that John is a doctor, which is a theme, we would also know that when he treats a patient his goal is to cure him, as a plan for such a cure he may send the patient to a laboratory for some tests. We also have a script of a prototypical visit to a doctor. In addition to that, we also know that doctors are usually intelligent and that that many of them are friendly to their patients.

When people are dealing with simple physical events such as going, moving objects, eating, etc. they use a different conceptual vocabulary than when they deal with social or political events, or when they talk about nuclear physics. In a social or political domain people think in terms of authority and power relations which can be represented as special memory structures called <u>triangles</u> (Schank and Carbonell 1978), rather than in terms of such primitive acts as moving a body part (MOVE) or grasping an object (GRASP).

The above concepts are used extensively in this thesis and I will assume that the reader is familiar with them.

5.2. Lexical Knowledge

5.2.0. Dictionary organization

In order to understand natural language we have to know the meaning of words and how the words are used. This information is contained in the parser's dictionaries. In this section we discuss the organization of these dictionaries.

A word can have many senses. The word "bank", for example, has at least two senses: the financial institution and the ground bordering a river. The simplest way to organize a dictionary is to list under each word all its senses. In this schema, when the parser reads a word it goes through the list of its senses looking for the one appropriate in the current context. This search can be made more efficient by sorting word senses according to the frequency of their use.

This simple ordering, however, is not a very good model of dictionary organization. For example, if we ask a person who is watching a baseball game to list all the senses of the word "base", he probably would mention the "baseball" sense of this word first. However, a person on a military base would probably mention the "military base" sense first, and might even not mention the "baseball" sense of at all if this person does not know much about sports.

This example suggests that a dictionary like Webster's where all the senses of a word are listed under one entry in the order of usage assuming a very "neutral" context is not a good model of how people organize their lexical knowledge. A better model would be to

attach a specialized dictionary to every stereotypical situation which requires a specialized vocabulary. When this situation is recognized as the current context, the parser should first consult this specialized dictionary while looking for a word. The following diagram illustrates this approach.

```
-------------------------------------------------------

   Baseball Dictionary:

   WORD: BASE    SENSES:  (1) S1

-------------------------------------------------------

   Military Dictionary:

   WORD: BASE    SENSES:  (1) S2

-------------------------------------------------------

   Main Dictionary:

   WORD: BASE    SENSES:  (1) S1  (2) S2

-------------------------------------------------------

          |-----------|       |-----------|
Word      |           |       |           |
          | "Baseball |       | "Military |
senses:   |  Base"    |       |  Base"    |
          |           |       |           |
          |-----------|       |-----------|

               S1                  S2

-------------------------------------------------------
```

Figure 5.1: Dictionary Organization.

When the parser reads the word "base" in a baseball context, it first looks it up in the current specialized dictionaries (the baseball dictionary, in this case). Specialized dictionaries contain pointers to only those word senses which are peculiar to the situation which activated this dictionary. If the word is not found in the specialized dictionaries or the senses listed in these dictionaries do not apply, the parser looks up the word in the main dictionary. In the present implementation specialized dictionaries are associated with scripts and are activated and deactivated following the activation and deactivation of the corresponding scripts.

Let us now concentrate on a single-sense entry. This entry consists of a conceptualization built from primitive concepts with various specifications and restrictions placed on their slots. We distinguish three types of specifications that particular words of a given natural language place on their underlying concepts:

1. Definitional Specifications,

2. Restrictions and Preferences,

3. Filler Expectations.

We discuss these specifications in the following three subsections.

5.2.1. Definitional Specifications.

The meaning of actual natural language words is usually represented by complex conceptualizations built out of primitive concepts. For example, there is no primitive act which can adequately represent the meaning of the verb "to say". Its meaning can be represented only by the combination of the primitive act MTRANS with the primitive act SPEAK. The simplest case of such combinations is when the definition of a word specifies the fillers of some slots of the concept underlying this word. We call this type of definitional specification slot fixing. Slot fixing is used to differentiate the meanings of words referring to the same conceptual frames. The difference between the verbs "to cry" and "to sweat" is in the fillers of the OBJECT and FROM slots of the EXPEL frame underlying both verbs.

(5.2) TO CRY:

```
(EXPEL ACTOR  X
       OBJECT (*WATER*)
       FROM   (*EYES* PART-OF X))
```

(5.3) TO SWEAT:

```
(EXPEL ACTOR  X
       OBJECT (*WATER*)
       FROM   (*SKIN* PART-OF X))
```

The most common use of definitional specification is in the differentiation of actions according to their instrumentality. Verbs like "to drive", "to fly", "to swim", etc. specify the instrumental aspect of a PTRANS. For example, the filler of the INSTRument slot of the concept underlying the verb "to drive" is the $DRIVE script with the ACTOR of the PTRANS playing the role of the DRIVER in it.

Another interesting example of definitional specifications is what we call slot tying. The English verbs "to go", "to walk", "to run", etc. specify that the filler of the OBJECT slot of the PTRANS frame underlying these verbs must be the same as the filler of the ACTOR slot. A different verb can put different restrictions on the same conceptual frame. For example, the verb "put" in "John put the book on the table" normally requires different objects in the ACTOR and OBJECT slots of the PTRANS frame.

Slot tying is used very extensively when the concept underlying a word involves more than one primitive conceptualization to represent its meaning. The verb "to buy", for example, represents two transfers of ownership: the merchandise from the seller to the buyer and the money from the buyer to the seller. This requires two ATRANS's:

```
(5.4) (AND
        EL (ATRANS ACTOR Z OBJECT Y       FROM Z TO X)
        EL (ATRANS ACTOR X OBJECT (MONEY) FROM X TO Z)),
```

where X is the buyer, Z is the seller, and Y is the object of the sale. The verb "to buy" ties the slots of the two ATRANS's which represent its meaning so that the ACTOR of the first is identical to the filler of the FROM slot of the same ATRANS and the filler of the TO slot of the second ATRANS. The same thing is true for the second ATRANS.

5.2.2. Restrictions and Preferences.

What is the difference between the meanings of the verbs "to drink" and "to eat?" Both of these verbs refer to an INGEST:

```
(5.5) (INGEST ACTOR X OBJECT Y TO (MOUTH PART-OF X))
```

However, the first verb implies that the OBJECT of INGEST must be a liquid without solid components. The second implies that the OBJECT contains solid components. These specifications do not fix the OBJECT slot to a specific conceptualization. Rather, they restrict the kind of concepts that can fill particular slots in a conceptualization by refining the requirements for the fillers of these slots.

Another example of slot restriction is the verb "to pour". This verb builds the following PTRANS frame:

```
(5.6)  (PTRANS ACTOR   X
               OBJECT  Y
               FROM    (INSIDE OBJ Z)
               TO      W)
```

In the general PTRANS frame, the filler of the OBJECT slot is expected to be a physical object, the filler of the FROM slot --

some location. The verb "to pour" further specifies that the filler of the OBJECT slot is supposed to be liquid-like and the filler of the FROM slot -- a location which is the inside part of some object. In general, there are many verbs specifically used when the object of PTRANS is liquid-like: to flow, to drip, to stream, to shower, to drizzle, etc.

The verb "to wheel" implies that the filler of the OBJECT slot of the corresponding PTRANS frame is expected to be a physical object with wheels but usually without an engine for self-propulsion. It also implies that the object is supported by the wheels while moving. The verb "to lift" indicates that the location in the TO slot must be higher than the location filling the FROM slot. We cannot say, for example, that John lifted a desk from the fifth floor to the fourth. The verb "to lower" has the opposite expectations. The verb "to land" in general means that the object was PTRANSed from some location above the ground to some location on the ground. The verb "to sink" means a PTRANS to some location under the water below the FROM location.

Most of the restrictions that words put on the slots of the underlying conceptualizations are preferential in the sense that some concepts are "better" fillers than the other. "Tea", for example, is a much better candidate for the filler of the OBJECT slot of the INGEST built by the verb "to drink" than "soup", although even the latter can be forced into the OBJECT slot. The OBJECT slot of the ATRANS underlying the verb "to give" is normally filled with a physical object, although an animal would also be quite acceptable. It is more difficult to force a human into this slot, although even this is possible, for example, especially if we are talking about slavery as in "The plantation owner gave two men to his son as a gift". (For another discussion of preference restrictions, see Wilks [1973]).

5.2.3. Filler Expectations.

So far we discussed only how words are related to underlying conceptual frames. We did not say anything about the relations between the words in a sentence. The most important aspect of these relations is the dependency between the concepts underlying these words. All the concepts used in a sentence must fit into some conceptual frame. If they do, we can figure out the meaning of a sentence even when what we hear is not a "grammatical" sentence.

This, however, does not mean that syntax does not play any role in language understanding. Words not only build conceptual frames with semantic requirements for every slot in them, but also often suggest how the fillers of these slots can be expressed in the surface sentence. Consider the following example:

(5.7) JOHN GAVE A BOOK TO MARY.

The word "gave" in the above sentence builds the following ATRANS frame:

```
(5.8) (ATRANS ACTOR   X   OBJECT Y
              FROM    X   TO     Z)
```

where the fillers of the ACTOR and TO slots are expected to be human and the filler of the OBJECT slot is expected to be a physical object. In addition to these expectations, the syntactic expectation for the ACTOR slot is: "the first PP before the verb". This is a fact about English and the word "gave", and not about the concept ATRANS. The filler of the OBJECT slot is expected to be expressed by a noun coming after the verb in the sentence with no prepositions in front of it. Syntactic expectations for the filler of the TO slot of the ATRANS built by the English word "gave" are relatively complex. The filler has to be expressed by a noun and is expected to come after the verb but before the filler of the OBJECT slot. In this case, no prepositions should be in front of it. If it comes after the filler of the OBJECT slot, it has to be preceded by the preposition "to".

It is interesting to compare syntactic expectations added by the word "gave" to the underlying ATRANS frame with the syntactic expectations added to the same conceptual frame by the Russian word "dal" (gave). The filler of the ACTOR slot of the ATRANS frame built by this Russian verb would normally be a noun in the nominative case, which is indicated by a specific ending of the noun. The filler of the OBJECT slot is expected to be expressed by a noun in the accusative case and the filler of the TO slot by a noun in the dative case. The word order in Russian is unimportant. In fact, in Russian we can use the four words: "John", "book", "Mary", "gave" in any order and the sentence will still be considered normal as long as the nouns are in the appropriate cases. Thus, cases play the same role in Russian as word order and prepositions in English, i.e., they are indicators used to locate fillers of conceptual frames.

5.3. Linguistic knowledge.

In the previous section we introduced the notion of syntactic expectations which use numerous "clues and indicators" provided by the language to efficiently locate fillers of conceptual frames. These features of natural languages are usually studied by morphology and syntax. The following features are used by the parser most extensively:

1) Word order

2) Word adjacency

3) Prepositions

4) Part of speech classification

5) Verb endings

Word order is by far the most important feature of English. It is used in two different ways: in clause processing and in noun group processing. In clause processing the significant indicator is whether the noun comes before or after the verb. In most verbs (unless they are in passive) the actor of the underlying conceptualization comes before the verb. The fillers of other slots usually come after the verb.

Word order is crucial in noun group processing. This has to do with the inherently linear structure of English noun groups and the crucial role played by another important syntactic feature - word adjacency. Depending on the word order inside the noun group, the same words can be interpreted as belonging to a different part of speech thus leading to a different meaning of the combination. For example, the phrase "North American" does not mean the same thing as the phrase "American North".

Combined expressions inside a noun group (i.e., those expressions whose meaning is determined by a combination of more than one word) are formed only from adjacent words. The insertion of an extra word inside a noun group usually leads to a different interpretation. For example, the phrase "dog owner" changes its meaning after the insertion of the word "shop": "dog shop owner". The same adjacency principle prevents us from interpreting the phrase "old man's car" as "man's old car", which means a different thing. Adjacency is also used in relative subclause, prepositional and participial phrase attachment which usually immediately follow the nouns that they modify.

The third syntactic indicator in natural language extensively used by our analyzers is prepositions. Prepositions mark concepts underlying the words that they precede thus guiding the process of slot filling. In the previous chapters we discussed the parser's handling of prepositions in general. In the next three chapters we give a detailed discussion of the preposition "by".

The fourth major feature of natural languages used by our analyzers is the differentiation of words into syntactico-morphological classes such as <u>parts of speech</u>. Parts of speech are used in two ways: as markers on the concepts in the slot filling process and, second, as indicators of various stages in the clause and noun group handling processes. For example, in a subclause the second verb normally indicates the end of the subclause. Inside a noun group an adjective normally cannot follow a noun, a number cannot follow either a noun or an adjective, etc. If such thing occurs it is normally an indicator of the end of the group.

Verb endings are used to recognize tenses and passive, progressive, and participial constructions. When the parser reads a word it first checks if it has it in its dictionaries. If it does not, it checks the list of irregular verb forms. If the parser does

not find the word on this list, it tries to find the root by stripping standard endings such as -ed, -ing, etc. Then the parser checks again whether the root is in the dictionaries.

How is linguistic knowledge used by the parser? There are three places in the parsing process where it is used: (1) in the syntactic expectations attached to the slots of conceptual frames, (2) in the control structure of the noun group processor, and (3) in the experts which process unpredicted input. We do not keep linguistic knowledge in a separate module. Instead, it is distributed throughout the whole parser and applied only when needed.

5.4. Contextual Knowledge.

Natural language understanding is done in context and is possible only in context. In this section I will illustrate this point with a few examples, and outline the most important areas in language analysis where context application is most crucial. Detailed discussions of particular mechanisms for handling context application will be contained in the further chapters.

The first major area of context application is word sense disambiguation. Many words have different meanings in different contexts. Consider the following sentence:

(5.9) JOHN GAVE MARY A TIP.

Depending on the context this sentence can mean "John gave Mary some money" or "John told Mary something helpful".

Even when the meaning of every word in a sentence is unambiguous, the sentence can have different meanings in different contexts. This happens most often in sentences with prepositional or participial phrases. For example, in the following sentence:

(5.10) JOHN SAW MARY DRIVING DOWN THE STREET.

the answer to the question "who was driving?" depends entirely on the context. By manipulating the context we can cause a participial or prepositional phrase to be attached to a different PP. For example, in

(5.11) MOHAMMED ALI WAS ATTACKED. HE HAD A BASEBALL BAT.
ALI HIT THE ATTACKER WITH THE BAT.

we interpret the phrase WITH THE BAT to be instrumental to ALI's action since we know from our world knowledge that bats are used in hitting things and if someone has a bat and is being attacked, he is likely to use the bat. However, the appropriate context can override this expectation:

(5.12) MOHAMMED ALI WAS ATTACKED BY TWO MEN. ONE OF THEM WAS
ARMED WITH A BASEBALL BAT, THE OTHER WITH A CHAIN.
ALI HIT THE MAN WITH THE BAT.

In this case, we conclude that Ali chose the man armed with the bat and used his fists for hitting. What happened here was that by the time we reached the last sentence in the above example, the focus of attention switched from Ali to the attackers. We were expecting to hear whom the famous boxer would punch first and, since the attackers were identified by their weapons, we were expecting to hear about the weapon of the attacker who was to be punched.

When we have more than one action in one sentence and an instrumental construction which can be interpreted as an instrument to both, only the context and world knowledge can point to a correct interpretation. Consider:

(5.13) THE USSR STRONGLY REACTS TO THE UNITED
STATES' DESTABILIZING THE BALANCE OF POWER
BY NORMALIZING RELATIONS WITH CHINA.

Who normalized relations with China, the USA or the USSR? The sentence itself does not give an answer. The instrumental phrase builds a conceptualization with an empty ACTOR slot. In order to choose between the two candidates to fill this slot we have to to know the political situation at the time.

Pronominal reference disambiguation is another area where context is crucial for understanding. The problem of how to determine who or what is referred to by such words as "he", "she", etc. is one of the most difficult and widely discussed in natural language (see, for example, Charniak [1972]). Syntactic theorists showed a lot of ingenuity trying to find formal syntactic rules which would locate the co-referent of a pronominal in the same sentence (see, for example, Ross [1967]). They did not have much success, but even if they did they still would have had to demonstrate that their rules could be used in the process of language understanding. Here, I will give a few example where the context plays a decisive role in the determination of pronominal referents.

Often, the referent to a pronominal is not mentioned in the same sentence. Consider:

(5.14) JOHN ASKED MARY FOR A BOOK.
SHE GAVE IT TO HIM.

All pronominals in the second sentence refer to the entities introduced in the first and this is a rather simple example of the problem. In a telephone conversation a pronominal can refer to something mentioned in a previous conversation two weeks ago or to something which was never mentioned but easily understood by both

participants. For example, in the case of an important announcement in a newspaper, a person can call his friend and say: "Did you see it?!" Even in a simple story, the referent of a pronominal can actually come after the pronominal itself as it is in the following example:

(5.15) JOHN SAID: "IT COSTS $4000."
HE POINTED AT HIS NEW CAR.

There is, however, an even more serious problem associated with pronominal reference specification. Many verbs, including such a common one as "to give", require the knowledge of the object for their disambiguation. For example, "to give a book" is an ATRANS, but "to give a kiss" is something quite different. Consider the following examples:

(5.16) JOHN ASKED MARY FOR A CANDY.
SHE GAVE HIM ONE.

and

(5.17) JOHN ASKED MARY FOR A KISS.
SHE GAVE HIM ONE.

The meaning of the second sentence in the above examples depends on the referent of the word ONE. We have to know the referent in order to choose the correct meaning of the verb "to give".

Another example where the context determines the meaning of the communication is sentences with such words as "yes" and "no". Consider the following story:

(5.18) JOHN ASKED MARY IF SHE HAD A PEN.
MARY SAID YES.

The meaning of the second sentence is a communication by the way speaking of the fact that Mary has a pen. The meaning of the word "yes" is determined entirely by the context. In our example it is to find the request, to determine the type of question and the desired answer and only then to build the appropriate representation. In fact, in order to understand the last example, the listener has to simulate the whole process of question answering.

The examples in this section demonstrate that context is a very important source of information in natural language understanding. We outlined several areas where context application is most crucial. These areas are: (1) word sense disambiguation, (2) prepositional phrase attachment, (3) participial phrase attachment, (4) relative subclause attachment, (5) pronominal reference specification, and (6) determination of the meaning of sentences containing such words as "yes" and "no".

5.5. Summary

In this chapter we described four general sources of knowledge used in language understanding: lexical knowledge, static world knowledge, linguistic knowledge, and contextual knowledge. Most of the parser's knowledge is stored in its dictionaries. Lexical knowledge, linguistic knowledge, and a great deal of static world knowledge can be stored in this way in the form of various requirements placed on conceptual frames and expectations attached to their slots. The parsing mechanisms described in the first part of the thesis achieve a considerable degree of integration in the application of these types of knowledge. The weak side of these mechanisms is their inadequate handling of contextual knowledge which, as the examples in the last section show, is crucial for natural language understanding. There is only one way in which the context can actively direct the current parser. This is done by using specialized dictionaries providing preferred meanings to the words specific to a given context. Consider the following example:

(5.19) JOHN ORDERED A SUB TO GO.

The preferred meaning of the verb "to order" in a fast-food restaurant context is "to order food" which builds a conceptual frame with two expectations: one for food and the other for the type of the order. The first of these expectations disambiguate A SUB and the second interprets TO GO. If, on the other hand, the same sentence was used in the context of a submarine base, the preferred meaning of the words "to order" and "sub" would have been quite different. This mechanism, however, is insufficient for handling most of the examples in the last section which require more direct predictive involvement of the context into the parsing process. In the next three chapters we explore the limits of the mechanisms described in the first part of the thesis, demonstrate their strong points in handling the first three types of knowledge and make explicit their shortcomings in the application of contextual knowledge.

CHAPTER 6

THE PASSIVE CONSTRUCTION

6.0. Introduction

The preposition "by" is interesting because of the variety of roles it can play in a sentence. The range of its use extends from subject designation in passive constructions to the indication of relevant PLAN-BOXES in instrumental constructions. The correct interpretation of this preposition requires application of many kinds of knowledge from syntactic to knowledge about scripts, plans, people's social roles, physical environment, etc. This richness of the preposition "by" made us chose it as a guide in our walk through various parsing problems. In this chapter we will discuss the passive construction.

Passive constructions account for the great majority of cases in which the preposition "by" is used. The front pages of two randomly chosen issues of The New York Times contained 22 instances of this preposition. In 12 cases the preposition "by" was used in a passive construction. Passive constructions are interesting for two reasons: first, their analysis requires fine interaction between syntactic and semantic knowledge, and, second, they require a different kind of request interaction from the ones introduced earlier, but one which, nevertheless, can be handled in the same parsing framework.

How different is the processing of passive constructions from the corresponding active constructions? Consider the following examples:

(6.1) JOHN WAS HIT BY MARY.

(6.2) MARY HIT JOHN.

In both sentences, the verb "hit" builds a PROPEL frame:

(6.3) (PROPEL ACTOR X OBJECT Z FROM X TO Z)

When the verb "hit" is used in the active construction, the filler of the ACTOR slot is expected to be the concept of a HUMAN underlying the first noun before the verb(*). Thus, in addition to semantic requirements for the filler of the ACTOR slot, this expectation uses such syntactic features as part of speech classification and word order. The filler of the OBJECT slot normally underlies the direct object of the verb. In the passive construction, however, expectations associated with the ACTOR and the OBJECT slots use different syntactic markers. The filler of the ACTOR slot is expected to come after the verb and be preceded by the preposition "by". The filler of the OBJECT slot is the first noun of the sentence. Semantic requirements for the OBJECT and the ACTOR slots remain unchanged.

The question now is whether we should have two different definitions of the verb "to hit", one for active and one for passive, or we should use only one dictionary definition and introduce special rules which would "fix" the standard definition for active and passive. These rules should be independent of individual verbs and work on most of them. The first approach was taken in the earlier versions of ELI where each verb had two definitions, one for active and one for passive. The second approach is more efficient and, intuitively, seems to be more plausible psychologically. Thus, in the later versions of the analyzer, we took the second approach. We will first sketch how this approach works and then describe it in greater detail.

6.1. Simple passivization rules.

We handle automatic passivization by storing only the "active" definition of the verb and modifying it when a passive construction is recognized. In a passive construction, the verb "to be" is typically followed by the past participle of the main verb. Thus, in the parser, the verb "to be" sets up an expectation for a past participle. English does not normally allow any noun groups or verbs between the verb "to be" and the past participle. If the expectation which looks for a past participle is satisfied, syntactic expectations attached to the slots of the conceptual frame underlying the main verb are rearranged according to the passivization rules stored in the definition of the verb. Normally, these are standard rules. I will first explain with a simple example what we want these rules to do, after that I will give a simplified version of these rules, explain how they work, and, then, gradually introduce more and more sophistication into them in order to handle more complex examples. Let us consider example (6.1) again:

(*) We call the first noun group before the verb the <u>subject</u>, and any noun groups not preceded by prepositions and which come after the verb the <u>direct object</u>. We will call all other noun groups (except appositives) <u>indirect objects</u>. (This does not necessarily correspond to the usual linguistic terminology.)

(6.1) JOHN WAS HIT BY MARY.

As we said earlier, the verb "hit" builds a PROPEL frame:

(6.3) (PROPEL ACTOR X OBJECT Z FROM X TO Z)

where the ACTOR slot is expected to be filled by the concept of a HUMAN in the subject position in the active sentence, and the OBJECT slot is expected to be filled by a physical object or an animate in the direct object position. Essentially, what the passivization rules are supposed to do is to change the syntactic parts of the above expectations to look at the "by" phrase instead of the subject and the subject instead of the direct object. Thus,

PASSIVIZATION RULES:

1) Expectations which look for concepts in the subject position are changed to look for the same concepts in the indirect object position preceded by the preposition "by".

2) Expectations which look for concepts in the direct object position are changed to look for the same concepts in the subject position.

Before I explain how these rules work, I would like to emphasize the point that the above rules are used to save memory. They allow us to store only one set of syntactic expectations associated with the conceptual frame of each verb instead of two: one for active and one for passive. However, there are many verbs to which the standard rules do not apply and whose definition indeed contains more than one set of expectations. The verb "to go" is an example of such a verb. The phrase "John is gone" is not a passive construction.

The next example illustrates how the standard passivization rules work. We will use the sentence "John was hit by Mary" from example (6.1).

1) The analyzer reads in the word "John". Since there are no active expectations at this moment, the analyzer stores the concept underlying the word "John" in its short term memory.

2) The analyzer reads in the word "was". This word does not build any structure but sets up several expectations designed to disambiguate its meaning. Among these expectations there is one which looks for a past participle. If this expectation is satisfied, the passive situation will be recognized and the passive construction "specialist" will be triggered.

3) The analyzer reads in the word "hit". This word is ambiguous even syntactically. It can express the present and past tenses

and the past participle of the verb "to hit". But only one of these three possibilities is expected by the expectations set up at stage 2. Thus, the analyzer chooses the "past participle" interpretation, triggering the passive construction "specialist". The definition of the verb "to hit" specifies that this verb has a standard passive form. The only slot in the PROPEL frame (6.3) which looks at the NP's before the verb is the ACTOR slot. The passive specialist changes its syntactic expectation according to Rule 1 to look for the preposition "by". The only slot which looks for an NP after the verb and without prepositions is the OBJECT. Its syntactic expectation is changed to look at the NP before the verb (Rule 2). The first NP is "John". "John" satisfies the semantic requirements for the OBJECT slot. Hence, the OBJECT slot gets filled.

4) Next the analyzer reads in the word "by". This word does not build any structure on its own. It is left in the short-term memory as a syntactic marker.

5) The last word of sentence (6.1) is "Mary". It is a noun group which is marked by the preposition "by". Therefore, the expectation attached to the ACTOR slot gets to examine it. Mary satisfies the requirements for the ACTOR slot of the PROPEL and the ACTOR slot gets filled.

The above standard passivization rules are too simplistic on several counts. First, there are cases when a verb can be followed by more than one direct object. For example, the verb "to give" in the active construction builds an ATRANS frame:

(6.4) (ATRANS ACTOR X OBJECT Y FROM X TO Z)

in which both the OBJECT and the TO slots can be filled by the concepts expressed in English by the noun phrases without prepositions. For example, in "John gave Mary a book", both "a book" and "Mary" are noun groups without prepositions. Mary fills the TO slot and "a book" fills the OBJECT slot. Correspondingly, the verb "to be" has two possible passive constructions:

(6.5) MARY WAS GIVEN A BOOK BY JOHN.

and

(6.6) A BOOK WAS GIVEN TO MARY BY JOHN.

with the filler of the TO slot appearing in the beginning of the sentence in the first example and the filler of the OBJECT slot in the second. Our standard passivization rules do not tell us which expectation should be redirected to look at the first noun group. If we redirect both of them, then we will miss the other filler which comes after the verb. The solution to this problem is not to redirect the expectations looking for direct objects, but make these expectations look in both places: before the verb and after. The following is the revised version of passivization rule 2:

PASSIVIZATION RULE 2a:

> For every request which looks for concepts in direct object positions a copy is made whose syntactic expectation is changed to look at the concept in the subject position.

Let us examine in greater detail how this works. Below are the slots of (6.4) with their expectations. (Expectations are parts of the requests which are managed by the parser's control mechanisms).

=== ACTOR SLOT ===

REQUEST1:

EXPEC: Semantic: Animate (or anything that can act like one)

Syntactic: Subject

ACTION: Fill the slot.

=== OBJECT SLOT ===

REQUEST2:

EXPEC: Semantic: Non-human object

Syntactic: Direct object

ACTION: Fill the OBJECT slot with the current concept. If it is animate, put REQUEST3 on the CONTINGENCY line.

REQUEST3: (on the CONTINGENCY LINE)

EXPEC: Semantic: Physical object and the OBJECT slot has already been filled with an animate but the TO slot is empty.

Syntactic: Direct object

ACTION: Put the filler of the OBJECT slot into the TO slot. Fill the OBJECT slot.

=== TO SLOT ===

REQUEST4:

EXPEC: Semantic: Human (or something that can act like one)

Syntactic: Direct object.

ACTION: Fill the TO slot. If the OBJECT slot has not been filled, put REQUEST5 on the CONTINGENCY line.

REQUEST5:

EXPEC: Semantic: Animate (or anything that can act like one)

Syntactic: Indirect object with preposition "to".

ACTION: Fill the TO slot; if the TO slot has already been filled, put its present filler into the OBJECT slot and fill the TO slot with the new concept.

Normally, when we hear "X gave Y...", where Y is not a person or an organization, we assume that Y is the object of giving, as in: "John gave a book to Mary". However, when Y is an animate object (an animal, for example), there is a possibility that this assumption is wrong as in sentence (6.8) below. As we are about to see, placing REQUEST3 on the CONTINGENCY line of requests, which is checked only if something goes wrong, takes care of this problem.

Let us, first, follow the interaction between these requests in the active case. Consider the following three sentences:

(6.7) MARY GAVE THE DOG TO JOHN.

(6.8) MARY GAVE THE DOG A BONE

(6.9) MARY GAVE JOHN TO THE DOG.

The state of the analyzer after it reads the words MARY GAVE can be described as follows:

```
|-----------------------------------------------|
| TOP CONCEPT: (ATRANS ACTOR (MARY) OBJECT Y    |
|                      FROM  (MARY) TO     Z)   |
| ACTIVE LINE:      REQUEST2, REQUEST4, REQUEST5|
| CONTINGENCY LINE: EMPTY                       |
|-----------------------------------------------|
```

Figure 6.1: STATE1

REQUEST1 is not on the ACTIVE line of requests because it has been triggered by MARY. In sentences (6.7) and (6.8), the next input is THE DOG which is not a human. Thus, REQUEST2 is triggered; it fills the OBJECT slot of the TOP concept. Since dogs are animate, REQUEST2 puts REQUEST3 on the CONTINGENCY line. Thus, the state of the analyzer after processing MARY GAVE THE DOG... is as follows:

```
|---------------------------------------------------|
| TOP CONCEPT: (ATRANS ACTOR (MARY) OBJECT (*DOG*)  |
|                      FROM  (MARY) TO      Z)      |
| ACTIVE LINE:       REQUEST4, REQUEST5             |
| CONTINGENCY LINE: REQUEST3                        |
|---------------------------------------------------|
```

Figure 6.2: STATE2

The next and the last phrase in sentence (6.7) is TO JOHN. REQUEST5 is triggered; it fills the TO slot of the TOP concept which completes the analysis of this sentence. In sentence (6.8), on the other hand, the next word is A BONE. All the requests on the ACTIVE line fail. Hence, the control "traps" to the CONTINGENCY line where REQUEST3 is successfully executed.

Let us now return to sentence (6.9). When the analyzer reads the word JOHN it is in STATE1. At this point, there are three active requests but only one of them, REQUEST4, looks for a human in the direct object case. Thus, REQUEST4 is triggered and it fills the TO slot of the TOP concept. The state of the analyzer is as follows:

```
|---------------------------------------------------|
| TOP CONCEPT: (ATRANS ACTOR (MARY) OBJECT Y        |
|                      FROM  (MARY) TO     (JOHN))  |
| ACTIVE LINE:       REQUEST2                       |
| CONTINGENCY LINE: REQUEST5                        |
|---------------------------------------------------|
```

Figure 6.3: STATE3

Normally, at this point, we would expect to hear the object of Mary's ATRANS to John as in "Mary gave John a dog". This would be handled by REQUEST2. However, something unusual might happen as in sentence (6.9), where it was John who was given to the dog. REQUEST2 expects a direct object, therefore it fails in this case. Hence, the process "traps" to the CONTINGENCY line where REQUEST5 corrects the situation.

Let us now examine the passive case and follow the effects of the passivization rules on the above expectations. Consider now the following sentence:

(6.10) A DOG WAS GIVEN A BONE BY MARY.

We will start at the point where the word GIVEN is read and the passivization rules are being applied.

Following the first passivization rule, REQUEST1 is modified as follows:

=== ACTOR SLOT ===

REQUEST1a:

EXPEC: Semantic: Animate (or anything that can act like one)

Syntactic: Indirect object with the preposition "by".

ACTION: Fill the ACTOR slot.

The following requests are added, according to the second passivization rule.

=== OBJECT SLOT ===

REQUEST2a:

EXPEC: Semantic: Non-human object

Syntactic: Subject

ACTION: Fill the OBJECT slot with the current concept. If it is animate, put REQUEST3 on the CONTINGENCY line. Remove REQUEST4a.

=== TO SLOT ===

REQUEST4a:

EXPEC: Semantic: Human (or something that can act like one)

Syntactic: Subject

ACTION: Fill the TO slot. If the OBJECT slot has not been filled, put REQUEST5 on the CONTINGENCY line. Remove REQUEST2a.

Note that there is no need to modify the CONTINGENCY requests. First, if they were to be modified, that should have been done at the time of their actual creation, e.g., at the time the verb is

read or even later. Second, the CONTINGENCY requests, by their very nature, look forward in the sentence for something unusual, but possible. The only thing that can come after the verb to which the passivization rules might apply is a direct object. When we apply the second passivization rule to a request which looks for a direct object, its syntactic expectation is redirected toward the grammatical subject of the sentence. But by the time of the possible creation of the CONTINGENCY requests the subject of the sentence has already been processed and, therefore, cannot possibly create an unexpected situation. We will illustrate this in our example when we get to the creation of REQUEST3.

The state of the analyzer, after the application of the passivization rules but before the testing of active requests, can be diagrammed as follows:

```
|--------------------------------------------------|
| TOP CONCEPT: (ATRANS ACTOR X OBJECT Y            |
|                          FROM  X TO      Z)      |
| SLIST:          (*DOG*)                          |
| ACTIVE LINE:          REQUEST1a, REQUEST2,  REQUEST2a,|
|                       REQUEST4,  REQUEST4a, REQUEST5  |
| CONTINGENCY LINE: EMPTY                          |
|--------------------------------------------------|
```

Figure 6.4: STATE4

Thus, we have now six active expectations embodied in the six ACTIVE line requests. Two of them look at the subject of the sentence: REQUEST2a and REQUEST4a. The only expectation which can be satisfied at this point is that of REQUEST2a. Thus, the concept (*DOG*) fills the OBJECT slot of the TOP concept and the "trap" REQUEST3 is created and placed on the CONTINGENCY line. REQUEST3 is created to recover if our filling of the OBJECT slot was wrong and some future input builds the "real" filler. Thus, there is no point in applying the passivization rules to it which would redirect it to the "past", not to the "future". REQUEST2 and REQUEST2a disappear, since they were attached to the slot which has just been filled, and REQUEST4a disappears because it looked at the subject and the subject has already been processed. This leaves the analyzer in the following state:

```
|-----------------------------------------------|
| TOP CONCEPT: (ATRANS ACTOR X OBJECT (*DOG*)   |
|                          FROM  X TO      Z)   |
| ACTIVE LINE:      REQUEST1a, REQUEST4, REQUEST5 |
| CONTINGENCY LINE: REQUEST3                    |
|-----------------------------------------------|
```

Figure 6.5: STATE5

Next, the analyzer processes the phrase A BONE. All active requests are tested and all of them fail, since the only request looking for a direct object is REQUEST4, but, semantically, it looks for a human, which a bone is not. Hence, REQUEST3 is tried and it succeeds, resulting in the following state of the analyzer:

```
|-----------------------------------------------|
| TOP CONCEPT: (ATRANS ACTOR X OBJECT (*BONE*) |
|                     FROM  X TO     (*DOG*))  |
| ACTIVE LINE:        REQUEST1a                 |
| CONTINGENCY LINE: EMPTY                       |
|-----------------------------------------------|
```

Figure 6.6: STATE6

The only request left at this point is REQUEST1a. When the analyzer processes the next phrase "by Mary", REQUEST1a is satisfied which completes the analysis of the sentence:

```
(6.11) (ATRANS ACTOR (MARY) OBJECT (*BONE*)
                FROM  (MARY) TO     (*DOG*))
```

6.2. Problems with passives.

One of the main reasons why people use passive constructions is the possibility of leaving the actor of the action unspecified. For example: "John was given a book", "The computer was sold", "Mary was taken to Boston", etc. In terms of conceptual frames and their expectations it means that the slot whose expectation is looking for a noun phrase preceded by the preposition "by" does not have to be filled with a concept from the same sentence. In other words, this expectation is optional and the request which embodies it should be placed on the OPTIONAL line of requests.

This requires a change in the first rule of standard passivization. We have to add that the new, modified expectation, which used to look at the subject of the sentence and now looks at the "by" prepositional phrase, becomes optional. On the other hand, the slot whose expectation now looks at the subject of the sentence becomes obligatory, even if it was optional in the active frame. For example, we can say both "John pushed" and "John pushed the door". However, we can only say "The door was pushed".

So far, we talked about the verbs which allow standard passivization rules. There are, however, many verbs which do not allow standard passivization. Intransitive verbs, for example, do not have passive forms. One cannot say "John cried" in the passive voice. An apparent exception is the verb "to go" which allows a construction which looks very much like passive, but which is not: "John was gone".

The second rule of standard passivization specifies that the only expectations that can be redirected in the passive to look at the subject are those which look for direct objects in the active construction. Some verbs, however, behave differently in this respect. For example, we can say

(6.12) JOHN WAS YELLED AT BY MARY.

or

(6.13) THE PAINTINGS WERE LOOKED AT BY THE COMMITTEE.

although, in the corresponding active constructions "John" and "the paintings" are not direct objects of their verbs.

One possible explanation of this behavior is to consider the preposition "at" in the above examples to be a part of the verb itself like a particle, which actually makes it a different verb. For example, the verbs "to turn" and "to turn on" are entirely different: they have different meanings and different syntactic behavior. In the sentence "John turned on the radio", the radio is a direct object of the verb "to turn on". We can also say "John turned the radio on". The rules of standard passivization applied to the verb "to turn on" produce a perfectly normal sentence "The radio was turned on by John". This, however, does not work for the verb "to yell". "At" is not a particle in the sentence "Mary yelled at John" and there is no separate verb "to yell at" different from the verb "to yell" either semantically or syntactically. A simple test shows that "John" is not a direct object in the last sentence. We cannot say "Mary yelled John at". Hence, our passivization rules cannot account for sentence (6.12).

We cannot solve the problem by introducing special rules about the preposition "at" either. This preposition by itself cannot account for sentences (6.12) and (6.13). We cannot say, for example, "The airport was left at John by Mary" meaning "Mary left John at the airport". Other prepositions can be used in the same way "at" was used in sentences (6.12) and (6.13). For example: "John was lied to by Mary". The conclusion, therefore, is that some verbs have their own idiosyncratic syntactic expectations in passive for which the standard passivization rules should not be applied.

6.3. Past participles as post-nominal modifiers.

A special case of the passive construction is the use of past participles as post-nominal modifiers. For example, the sentence:

(6.14) THE BOOK GIVEN TO MARY BY JOHN WAS INTERESTING.

is a short way of saying "The book which was given to Mary by John was interesting". The conceptualization built by the first noun phrase of (6.14) is:

```
(6.15) X = (#PHYSOBJ TYPE (BOOK) REL Y)

       Y = (ATRANS ACTOR (JOHN) OBJECT X
                   FROM  (JOHN) TO     (MARY))
```

The second part of the above conceptualization is exactly the same as the one that would have been built by "The book was given to Mary by John". In fact, the same expectations can be used to fill the gaps of the above frame. Here, we will expand and elaborate the approach outlined in Section (4.4) on participial phrases.

We can imagine the following simple procedure for handling this kind of post-nominal modifiers:

1) First, we recognize that we indeed have a post-nominal modifier expressed as a past participial phrase. This is not as easy as it was in the case of a regular passive construction where the verb "to be" gave us "advance warning" and enabled us to set up expectations for the past participial form of the verb. In the case of post-nominal modifiers, however, we do not have such expectations. When we say "The book...", we do not expect a restrictive clause. But the fact that something is unexpected can also be very informative. A verb in the past participial form is not expected at this point. This forces the analyzer to "trap" and apply its general knowledge about the structure of English sentences. This knowledge says that a past participial construction following a noun group can serve as a post-nominal modifier, in which case, the following is done:

2) The conceptual frame underlying the verb is attached to the noun group being modified using the REL link. (This link can be translated into English as "such that").

3) The standard passivization rules are applied to the syntactic expectations associated with the slots in the conceptual frame.

4) The resulting requests are applied to the input sentence.

One problem with this procedure is immediately apparent. The standard passivization rules redirect expectations looking for the direct object of the verb to look at the subject which we defined as the first noun group before the verb. This can be stretched to account for "John ate the food given to him by Mary", since "the food" is a noun group without a preposition which precedes the verb under consideration "to give". However, this will not work in the case of:

(6.16) JOHN WENT TO THE CAR GIVEN TO HIM BY MARY.

since "the car" is preceded by the preposition "to".

The solution to this problem is to use only the first standard passivization rule, which redirects the request looking for the subject to look for a "by" prepositional phrase and which still

works. Instead of using the second standard passivization rule we can do the following. When we reach Step 3 of the above algorithm, we already have the noun group and the frame of the conceptualization which modifies this noun group. Now, instead of trying to modify the syntactic expectations of this frame which look for a direct object, we can apply their semantic parts to the noun group being modified, disregarding their syntactic parts. For example, when the analyzer reads in the word "given" in (6.16) it has already built:

```
(6.17) (PTRANS ACTOR (JOHN) OBJECT (JOHN)
                   TO (LOCOF OBJECT (#PHYSOBJ TYPE (CAR))
```

Then the past participial modifier is recognized and the ATRANS frame is appended to (6.17) resulting in:

```
(6.18) (PTRANS ACTOR  (JOHN)
               OBJECT (JOHN)
               TO     (LOCOF OBJECT X))

         X = (#PHYSOBJ TYPE (CAR)
                       REL (ATRANS ACTOR Y OBJECT X
                                   FROM  Y TO    V)))
```

After this, we have only two empty gaps: Y and V. The expectation attached to the gap Y was modified by the first passivization rule to look for a noun group preceded by the preposition "by". The expectation attached to the V slot was not changed and still looks for a "to" prepositional phrase. (All semantic expectations, of course, remain the same.) These expectations pick up the remaining noun groups of the sentence in the usual manner.

6.4. Summary.

The meaning of a passive construction is often, but not always, the same as the meaning of the corresponding active phrase. The difference is, primarily, in the focus of attention in the sentence. When we say "John hit Mary" we focus attention on John. When we say "Mary was hit by John" we focus attention on Mary. We have already discussed the case where passive constructions are used when the speaker does not want to specify the actor of the action. There may be other reasons why people use passive constructions. Here, we are not trying to answer the question why people use them but rather to illustrate how they can be handled in our parsing framework and illuminate the role played in these constructions by the preposition "by". Our main goal was to explore the interaction between semantic and syntactic knowledge used in the processing of passives.

This chapter shows that passive constructions can be processed by the mechanisms developed in the previous chapters of this thesis. We described the actions of passivization rules which modify parsing requests. This imposes a relatively rigid format on the request

structure. They have to clearly indicate and separate their semantic and syntactic expectations, structures on which these expectations are focusing, and target slots which they are filling. Other modules of the program should be able to look at them and figure out what they are doing and modify them, if necessary. Finally, we showed the interaction between the participial phrase "expert" described in Chapter 4 and the passivization rules introduced in this chapter. We also have shown that the preposition "by" plays only an auxiliary role in the processing of passive constructions. It is used only as a syntactic "case" marker and the information stored under its dictionary definition is, in fact, never consulted. This information will be discussed in the next chapters.

CHAPTER 7

INSTRUMENTAL CONSTRUCTIONS

7.0. Introduction.

The next most common use of the preposition "by" is in instrumental constructions. Instrumental constructions specify the details of how the main action underlying the sentence was accomplished. For example, when we say:

(7.1) JOHN WENT TO NEW YORK BY TRAIN.

we specify that John accomplished his physical transfer to New York as follows: he went to the railroad station, got on the train, and the train with John on board went to New York. In other words, we apply our knowledge about trains and transportation scripts (Schank and Abelson 1977, Cullingford 1978) and apply it to the specification of John's action. When we say:

(7.2) JOHN ENTERED THE HOUSE BY BREAKING THE WINDOW.

we mean that John broke the window and then physically transferred himself through this window from the outside to the inside of the house.

Thus, in order to understand instrumental constructions we have to examine the lower-level actions of which the action underlying the main clause is composed. This ability to expand single events into their component actions plays an important role in knowledge organization. Every action can be expanded into an _episode_ or a sequence of lower level events of which this action was composed. For example, when we say "John went to England", we can expand his trip by describing how he went to the airport, got his ticket, took the airplane, etc. On the other hand, if we are interested only in the fact of John's going to England, we can refer to his trip as a simple PTRANS. Thus, episodes help us to organize information hierarchically, leaving only the most important events on the top level. Instrumental constructions are used to specify the internal structure of episodes in a concise and efficient way.

There are three ways in which instrumental constructions can specify the internal structure of an episode: through an action, an instrumental script, and a physical object used in one of the component actions. Let us briefly introduce these methods.

Sentence (7.2) gives us an example of episode expansion through a constituent action. "Entering a house" consists of two actions: making an opening and using it as an entrance. Thus, BREAKING THE WINDOW specifies the first part of the episode. We discuss this type of episode expansion in Section 7.1.

Sentence (7.1) gives us an example of episode expansion through an instrumental script. Our knowledge about physical transfers tells us that people can physically transfer themselves using transportation scripts. Travelling by train is such a script and, therefore, can be fit into the internal structure of John's PTRANS to New York. This type of episode expansion is discussed in Section 7.2.

Sentence (7.3):

(7.3) JOHN ATTACHED THE SIGN TO THE DOOR WITH A SCREW.

gives an example of episode specification through a physical object used in one of the constituent actions. There are many ways to attach a sign to a door. It can be done with glue, nails, screws, etc. Each of these methods involves different actions. In our case, John used a screw driver, and drove the screw through the sign into the door. In Section 7.3 we discuss how our knowledge about physical objects can be used in determining the actions in the episode.

In English, there are three types of instrumental constructions which roughly correspond to the three methods of episode expansion introduced above. Two of them involve the preposition "by" and one the preposition "with". One is when the preposition "by" is followed by a progressive verb-construction as in "John passed the exam by studying hard". The second construction is "by" followed by a noun referring to an instrumental script, as in "John got a letter by mail". The third construction is the preposition "with" followed by a noun group specifying a physical object. The correspondence between the above constructions and the types of episode expansion is not very strict. For example, we can say: "John went to New York by taking a train", or "John went to New York by train", or "John went to New York with a train" and all of these sentences mean essentially the same thing, namely that John used the TRAIN script to get to New York. In fact, we will show in Section 7.4 that other languages, such as Russian, for example, have even weaker correspondence between the surface instrumental structures and the types of episode expansion, which does not mean that the speakers do not use these types of expansion.

Parsing and attachment of instrumental constructions requires knowledge of the internal structure of events described in the main clause. In this chapter we examine and briefly characterize the

internal structure of events and discuss how it can be used in the interpretation of instrumental constructions. The following three sections of this chapter correspond to the three types of episode expansion.

7.1. Fitting an action into an episode.

What is the meaning of instrumental constructions? Let us consider Sentence (7.1) in greater detail:

(7.1) JOHN ENTERED THE HOUSE BY BREAKING THE WINDOW.

Breaking the window was instrumental for John's getting into the house. We can represent this in CD as follows:

```
(7.4)  (PTRANS ACTOR  (JOHN)
               OBJECT (JOHN)
               FROM   (OUTSIDE PART (HOUSE))
               TO     (INSIDE  PART (HOUSE))
               INSTR  . )
                      |
         |------------|
         V
     (CAUSE CONCEPT1 (DO ACTOR (JOHN))
            CONCEPT2 (STATE-CHANGE STATE-NAME (PHYSTATE)
                                   OBJECT     (WINDOW)
                                   VAL-FROM   (+10)
                                   VAL-TO     (-10)))
```

But does this simple linking of the two concepts with a link called INSTR really illuminate the meaning? What do we mean by saying that one action is instrumental to another? In order to answer these questions we have to examine in greater detail the structure of the actions involved. The meaning of the sentence:

(7.5) JOHN ENTERED THE HOUSE.

can be summarized in Conceptual Dependency as:

```
(7.6) (PTRANS ACTOR (JOHN)
              FROM  (OUTSIDE OBJECT (HOUSE))
              TO    (INSIDE  OBJECT (HOUSE)))
```

But this is only a summary of the actions involved in the act. We have more detailed knowledge about how people enter houses: First, they make an opening (normally, they open the door) and then they physically transfer themselves inside through this opening. In our first example, the window is a separator which separates the inside of the house from the outside. Breaking the window transforms it from the SEPARATOR into a CONNECTOR thus making an opening into the house (a detailed discussion of SEPARATORS and CONNECTORS is contained in (Lehnert 1977)). Thus, breaking the window corresponds

to the first step in the "entering into a house" procedure, i.e., making an opening or, using Lehnert's terminology, establishing a CONNECTOR.

We can give further evidence in support of this interpretation of sentence (7.1). Our knowledge of "entering" tells us that once a CONNECTOR is established, it will be used for the actual physical transfer. Thus, if we are right and the breaking of the window establishes a CONNECTOR required in the first part of "entering", then an inference would be made that John entered the house through the window. Indeed, this is the case, which we can demonstrate by trying to force a different inference:

(7.7) JOHN ENTERED THE HOUSE BY BREAKING THE WINDOW
AND THEN WALKING IN THROUGH THE DOOR.

The above sentence sounds odd because we expected John to enter *through the connector* that he had established. Thus, breaking the window in sentence (7.1) refers to the establishment of the connector in the ENTRANCE episode.

The use of the preposition "by" combined with a progressive verb-construction (BY+ING) is a syntactic clue that indicates that the action introduced by the prepositional phrase is a part of the episode underlying the main action. The problem is determining how the instrumental action fits into the episode. We do not have a complete answer to this question, only a few heuristics that people seem to use. One of the heuristics is to check whether the action in the instrumental construction is one of the major actions normally used to achieve the main act. This explanation was used in our second example in this chapter, "John entered the house by breaking the window". In order to enter an enclosed area, first, we have to have an opening and then PTRANS ourselves through this opening from the outside inside. Breaking the window achieves the first part of "entering".

The question of what actions are "major" and "normally used" is also very unclear. In the case of scripts we know the exact sequence of events and which actions are the most important. Ironically, we know scripts "too well" to "open" them up with instrumental constructions. For example, the sentence:

(7.8) JOHN ATE IN A RESTAURANT BY ORDERING FOOD.

sounds odd, although ordering is one of the most important scenes in the restaurant script.

Probably the most natural and easiest case of instrumental attachment is when the instrumental construction specifies one of the PLAN-BOXES achieving the intent of the main action. PLAN-BOXES and intentionality are discussed in detail in (Schank and Abelson 1977, Wilensky 1978). Consider the following sentence:

(7.9) JOHN GOT THE MONEY FROM BILL BY PULLING A GUN.

Pulling a gun constitutes a threat, one of the standard PLAN-BOXES to achieve possession of an object. Thus,

Instrumental constructions can specify one of the PLAN-BOXES achieving the intent of the main action.

To further illustrate this point, we will take all the PLAN-BOXES for the D-CONT (delta control) goal listed in (Schank and Abelson 1977) and show how all of them can referred to by the BY+ING instrumental construction.

Schank and Abelson (1977) define eight PLAN-BOXES that define D-CONT: ASK, INVOKE THEME, INFORM REASON, BARGAIN OBJECT, BARGAIN FAVOR, THREATEN, OVERPOWER, and STEAL. We will paraphrase one of their examples without further comments.

=== ASK PLAN-BOX ===

JOHN GOT ONE OF THE HARRY'S BOTTLES OF SCOTCH
BY SAYING "HOW ABOUT ONE OF THEM FOR ME?"

=== INVOKE THEME PLAN-BOX ===

JOHN GOT ONE OF THE HARRY'S BOTTLES OF SCOTCH
BY REMINDING HIM WHAT TRUE FRIENDS THEY HAD BEEN.

=== INFORM REASON PLAN-BOX ===

JOHN GOT ONE OF THE HARRY'S BOTTLES OF SCOTCH
BY TELLING HIM THAT THE BOSS WANTED IT.

=== BARGAIN OBJECT PLAN-BOX ===

JOHN GOT ONE OF THE HARRY'S BOTTLES OF SCOTCH
BY OFFERING HIM A BOTTLE OF MONTRACHET IN EXCHANGE.

=== BARGAIN FAVOR PLAN-BOX ===

JOHN GOT ONE OF THE HARRY'S BOTTLES OF SCOTCH
BY OFFERING TO CUT HIS FIREWOOD.

=== THREATEN PLAN-BOX ===

JOHN GOT ONE OF THE HARRY'S BOTTLES OF SCOTCH
BY MAKING HARRY AN OFFER HE COULDN'T REFUSE.

=== OVERPOWER PLAN-BOX ===

JOHN GOT ONE OF THE HARRY'S BOTTLES OF SCOTCH
BY PUNCHING HARRY IN THE NOSE.

=== STEAL PLAN-BOX ===

JOHN GOT ONE OF THE HARRY'S BOTTLES OF SCOTCH
BY QUIETLY PUTTING IT IN HIS BIG BRIEFCASE.

An interesting problem related to PLAN-BOXES and NAMED PLANS (Schank and Abelson 1977) arises when we try to explain why we can say:

(7.10) JOHN GOT INTO THE HOUSE BY TAKING THE KEY FROM MARY.

and would not normally use:

(7.11) JOHN ENTERED THE HOUSE BY TAKING THE KEY FROM MARY.

Somehow "entering" is a more limited episode than "getting into" and "getting the key" is a normal thing to do in order to "get into" a house, while it does not seem to belong to the act of "entering" itself. The difference lies in the fact that "getting into" implies a general D-PROX goal (achieve being inside) while "entering" implies a rather specific NAMED-PLAN which only includes "making an opening" and "physical transfer through this opening from the outside to inside". In general, the more restricted is the NAMED-PLAN underlying the main action, the fewer alternative PLAN-BOXES it contains, - the fewer is the choice of instrumental constructions. In fact, the most restricted NAMED-PLANS, i.e., SCRIPTS, often cannot "accept" any BY+ING instrumental constructions. For example, the verb "to run" refers to the script $RUN (Schank and Abelson 1977) and I cannot think of anything which can be used instead of "doing x" in the following passage:

(7.12) JOHN RAN INTO THE HOUSE BY "doing x".

Thus, we have the following two principles:

Clauses with more restricted underlying NAMED-PLANS allow fewer instrumental specifications.

and

Clauses referring to SCRIPTS do not accept any instrumental specifications

In the above examples, we took a "top-down" approach considering how the parser could interpret the BY+ING constructions on the basis of its general knowledge about people's plans and goals. We did not take into account the effects that a particular context might have on the interpretation of instrumental constructions. An appropriate context can make almost any BY+ING construction meaningful with almost any main clause. Consider the following example. Suppose we want to express the fact that John went to New York and that on the way he stopped in Milford for a cup of coffee. We cannot use:

(7.13) JOHN GOT TO NEW YORK BY STOPPING FOR A CUP OF COFFEE IN MILFORD.

although the coffee scene was part of his trip. "Taking a cup of coffee" does not satisfy any of the PLAN-BOXES normally associated with a D-PROX goal. Somehow the prepositional phrase has to be more "instrumental" in the sense that the listener has to be able to establish the enablement relation between this phrase and the main action. The establishment of this relation is accomplished in the process of understanding and involves both the knowledge of usual ways to achieve the main action and the particular context in which the sentence was used. For example, the sentence:

(7.14) JOHN GOT TO NEW YORK BY TAKING THE NINE O'CLOCK TRAIN.

sounds "normal" out of context since using a means of transportation is one of the possible plan-boxes to achieve a D-PROX which is presumably the goal of any PTRANS with identical ACTOR and OBJECT slots. However, we can make it sound "abnormal" by supplying the context in which John could not possibly take a train to New York, as in the following example:

(7.15) JOHN WAS IN LONDON. HE GOT TO NEW YORK BY TAKING THE NINE O'CLOCK TRAIN.

Context can make an "abnormal" sentence sound "normal":

(7.16) JOHN WAS FALLING ASLEEP AT THE WHEEL. HE GOT TO NEW YORK BY STOPPING FOR A CUP OF COFFEE IN MILFORD.

In the above passage, the first sentence makes us think of the enablement condition which we would not otherwise notice, namely that in order to do almost anything a person must be awake. Drinking coffee enables John to stay awake and, thus, to continue

his trip. These inferences could have been made even in the case of sentence (7.13) without any specific context, but normally people do not make them.

The above example suggests the following "bottom-up" heuristic for fitting an action into an episode. When the context sets up a violation of some important precondition we expect to see either the failure of the main action or some explanation of how the precondition was fixed. Thus, when we hear that John was sleepy in example (7.16), we note that this is a violation of an important precondition for a wide range of activities. When we hear that John got to New York and we assume that he was driving, we want to know how he did it. There are several normal ways of dealing with this problem. One is to take some other form of transportation, another is to drink coffee, a third is to have someone else drive instead of you, a fourth is to have someone talk to you while you are driving, etc. All of these explanations can be expressed by a BY+ING construction. Thus,

> Instrumental constructions can satisfy expectations set up by the violations of some preconditions of the main action.

In a sense, the above example brings us back to the PLAN-BOXES. In passage (7.16), when we hear that John was sleepy, we infer that his goal was not only to arrive in New York, but to arrive there safely. The above list of possible actions is nothing else than the list of PLAN-BOXES to achieve this goal. Thus, we can say that in this case too the instrumental construction had to specify one of the PLAN-BOXES for achieving the intended goal of the main action.

We do not claim that these heuristics solve the problem of how to "fit" an instrumental action into an episode. The goal of this discussion is to identify the problems and to point to the kind of knowledge needed for a solution. BY+ING constructions are used in English to introduce an action instrumental to the main action of the sentence. Instrumental actions either specify one of the ways to achieve the main action, or they satisfy some of the expectations set up by the context. A more specific answer requires more detailed study of people's goals and plans and the internal structure of episodes.

Let us now take an example of instrumental construction and see how it can be handled by the parser. We will use sentence (7.2) which we repeat here:

(7.2) JOHN ENTERED THE HOUSE BY BREAKING THE WINDOW.

First, the parser analyzes the phrase JOHN ENTERED THE HOUSE in the usual fashion. It produces the following conceptualization:

```
(7.17) (PTRANS ACTOR  (JOHN)
               OBJECT (JOHN)
               FROM   (OUTSIDE PART (HOUSE))
               TO     (INSIDE  PART (HOUSE))
               INSTR  . )
                        |
           |------------|
           V
        USE-DOOR, USE-OPENING
```

The instrument slot of the above conceptualization contains pointers to the usual ways of achieving this particular act. USE-DOOR is the most normal way of entering a house. It is a plan consisting of two elements: getting the door opened and going through it. It can be represented as follows:

```
(7.18) USE-DOOR:

    (GOAL TYPE  (ACHIEVE)
          STATE (DOOR CONNECTOR-ENABLE (PTRANS)))

    (PTRANS ACTOR  X   OBJECT X
            FROM       (OUTSIDE PART (HOUSE))
            TO         (INSIDE  PART (HOUSE))
            CONNECTOR  (DOOR))
```

where ACHIEVE represents a general goal of achieving a certain state of the world. ACHIEVE goals can suggest particular ways for their attainment depending on the state they are trying to achieve. In our example, for instance, we could have shown the instrument slot of the GOAL conceptualization containing pointers to such plans as USE-KEY, KNOCK-ON-THE-DOOR, etc.

USE-DOOR is a very specific plan for entering a house. In case this plan fails, we have another, more general plan for achieving the same goal:

```
(7.19) USE-OPENING:

    (GOAL TYPE  (ACHIEVE)
          STATE (Y CONNECTOR-ENABLE
                   (PTRANS ACTOR Z OBJECT Z
                           FROM      (OUTSIDE OBJ (HOUSE))
                           TO        (INSIDE  OBJ (HOUSE))
                           CONNECTOR Y)))

    (PTRANS ACTOR  X   OBJECT X
            FROM       (OUTSIDE PART (HOUSE))
            TO         (INSIDE  PART (HOUSE))
            CONNECTOR  Y )
```

where Y is an object which enables PTRANSs into the house (for

details about CONNECTORs see [Lehnert 1977]). Plans (7.18) and (7.19) are part of our knowledge about entering into buildings. We say that we have understood an instrumental construction used with JOHN ENTERED THE HOUSE when it achieves one of the components of these plans. Let us now return to the parsing process and see how this is done in our example.

When the parser reads the preposition BY followed by an ING form of the verb, it calls the BY+ING expert. This expert takes the concept underlying the subject of the sentence and puts it on the SLIST. Then it stacks the current TOP-CONCEPT, sets up an expectation for a verb in the ING form, and relinquishes control to the structure-driven mode of parsing. Thus the word BREAKING is interpreted as a verb and builds the following conceptualization:

```
(7.20) (CAUSE CONCEPT1 (DO ACTOR X)
              CONCEPT2 (STATE-CHANGE STATE-NAME (PHYSICAL)
                                     OBJECT     Y
                                     VAL-FROM   (+10)
                                     VAL-TO     (-10)))
```

with an expectation attached to the ACTOR slot looking for an animate object on the SLIST, and an expectation attached to the OBJECT slot looking for a physical object in the direct object position. JOHN satisfies the expectation of the ACTOR slot and the concept underlying the next input satisfies the expectation of the OBJECT slot. Thus, by the time the parser reaches PERIOD, it has built the following structure:

```
(7.21) (CAUSE CONCEPT1 (DO ACTOR (JOHN))
              CONCEPT2 (STATE-CHANGE STATE-NAME (PHYSICAL)
                                     OBJECT     (WINDOW)
                                     VAL-FROM   (+10)
                                     VAL-TO     (-10)))
```

The above conceptualization is not the only result of the participial phrase processing. When we parse the state change of the window in (7.21) the memory creates a token for this window (if it did not already have one) and modifies it according to the state change. Physical state is a rather abstract notion which means different thing for different objects. We have to have special inferences attached to every object which fire off whenever a change in the physical state of the object is asserted. For example, for a chair, physical state -10 means that one cannot sit on it, for a car it means that one cannot drive it, for a person it means that the person is dead. For a window, this state means that it enables PTRANSes. (This, of course, does not mean that "breaking" is the only way of making a window a CONNECTOR with respect to PTRANSes. Opening a window, for example, achieves this without changing the physical state of the window.) Thus, by the time the end of the participial phrase is reached, WINDOW used in (7.21) will look as follows:

```
(7.22) (WINDOW
        PHYS-STATE        (-10)
        CONNECTOR-ENABLE (PTRANS ACTOR Z OBJECT Z
                          FROM      (OUTSIDE OBJ (HOUSE))
                          TO        (INSIDE  OBJ (HOUSE))
                          CONNECTOR (WINDOW)))
```

At this point, control is returned to the BY+ING expert wich has to fit conceptualization (7.21) into one of the plans (7.18) and (7.19).

There are many ways in which actions can achieve goals (for a detailed discussion see [Schank and Abelson 1977] and [Wilensky 1978]). In our case, we have to check if (7.21) can achieve goals in (7.18) or (7.19). To do this we have to use a relatively general inference rule which says:

If a person has a goal of achieving a certain state of an object and he or she changes some state of this object, this might result in the achievement of the goal.

This rule is attached to the ACHIEVE type of goal and recognizes the fact that many properties of physical objects are related. Plan (7.18) is rejected right away because WINDOW does not match DOOR used in this plan. The parser tries plan (7.19). The above rule is triggered by the ACHIEVE goal and the concept WINDOW is checked as to whether it satisfies the goal in (7.19). The check succeeds and the plan is instantiated with conceptualization (7.21). The following is the final representation of "John entered the house by breaking the window":

```
(7.23) (PTRANS ACTOR  (JOHN)
               OBJECT (JOHN)
               FROM   (OUTSIDE OBJECT (HOUSE))
               TO     (INSIDE  OBJECT (HOUSE))
               INSTR  . )
    |---------------|
    V
   (GOAL TYPE   (ACHIEVE)
         STATE  (WINDOW CONNECTOR-ENABLE ...)
         INSTR
          (CAUSE CONCEPT1 (DO ACTOR (JOHN))
             CONCEPT2 (STATE-CHANGE STATE-NAME (PHYSICAL)
                                    OBJECT     (WINDOW)
                                    VAL-FROM   (+10)
                                    VAL-TO     (-10))))

       (PTRANS ACTOR      (JOHN)
               OBJECT     (JOHN)
               FROM       (OUTSIDE OBJECT (HOUSE))
               TO         (INSIDE  OBJECT (HOUSE))
               CONNECTOR  (WINDOW))
```

Thus, the parsing of even simple BY+ING instrumental constructions involves a great deal of knowledge about plans and goals, and properties of physical objects. The above example provides only a glimpse at a general solution to the problem of instrumental construction interpretation which requires much more developed representations for all this knowledge. However, our goal in this part of the thesis is not to provide exaustive solutions to every parsing problem. We have taken a breadth-first approach and attempted to point out the serious problems that have to be solved, while sketching possible solutions wherever it is possible.

7.2. Instrumental scripts.

The BY+ING construction is not the only way to introduce instrumentality in English. Instruments can be expressed by the preposition "by" followed by a word referring to an appropriate instrumental script (BY+INSTR constructions). For example, one can say:

(7.24) JOHN WENT TO BOSTON BY BUS.

In the above example, the phrase "by bus" refers to the "taking the bus" script. This script includes the roles for the passengers and the bus driver. It also specifies that, normally, passengers get tickets, get on the bus, give the tickets to the driver, sit down, the driver drives the bus, and, finally, they arrive and get off the bus. This script is only fleetingly referred to in sentence (7.24), but, if needed, it can be instantiated into the "John's trip to Boston" episode. For example, if sentence (7.24) is followed by:

(7.25) ON THE WAY, JOHN TALKED TO THE DRIVER.

we would open the bus script and infer that John was a passenger, that the driver was the driver of the bus, that John had a ticket and gave it to the driver, and that the driver was driving the bus to Boston at which point John talked to him.

We have the whole array of transportation scripts which can be used as instruments for a PTRANS. Usually, they are expressed by the preposition "by" followed by the name of the vehicle used in the appropriate script. Examples of frequently used transportation scripts include travelling by car, by train, by plane, by boat, etc.

Not all transportation scripts can be used with every PTRANS conceptualization. It would be odd to say:

(7.26) JOHN WENT FROM NEW YORK TO LONDON BY CAR.

or

(7.27) JOHN WENT FROM 5TH AVENUE TO 7TH AVENUE BY PLANE.

Our knowledge about PTRANS's and geography tells us what means of transportation would be possible and would be appropriate. Conversely, every transportation script specifies the necessary preconditions for its application. There has to be railroad service in order to go by train, there has to be a road in order to go by car, there have to be airports in order to go by plane, etc.

If instead of a well known transportation script in a BY+INSTR construction we use an unusual means of transportation people, would treat it by analogy with a transportation script using a similar vehicle. For example, the sentence:

(7.28) JOHN ARRIVED BY TRACTOR.

sounds a bit odd, but can be understood by analogy with "arriving by car".

Most of the instrumental scripts are expressed by a noun without an article which is preceded by the preposition "by". However, for historical reasons, some scripts use different constructions. For example, instead of saying:

(7.29) JOHN ARRIVED BY FOOT.

we use

(7.30) JOHN ARRIVED ON FOOT.

although, the former is perfectly understandable. Similarly, we do not normally say "John arrived by horse", although people would have no problems understanding it. There is no logical reason why we can say "John arrived by car" and cannot say "John arrived by horse".

Transportation scripts do not necessarily have to be specified through the name of the vehicle. We can say:

(7.31) JOHN LEFT BY AIR AND RETURNED BY SEA.

which is a grand way to say that John left by boat and return by plane. However, when we say:

(7.32) JOHN ARRIVED BY LAND.

we refer to the whole bundle of land transportation scripts.

Instrumental BY constructions are not limited to public or private transportation of people. If the OBJECT of the PTRANS is a physical object then an object transfer script can be used. For example:

(7.33) JOHN SENT MARY A BOOK BY MAIL.

or

(7.34) THE PACKAGE ARRIVED BY AIR FREIGHT.

Mail is an interesting script because it can serve as an instrument of not only a PTRANS but also a ATRANS, and an MTRANS. For example:

(7.35) JOHN BOUGHT A BOOK BY MAIL.

and

(7.36) JOHN NOTIFIED MARY BY MAIL.

The last two examples are different from the previous ones in that the script in the BY+INSTR construction does not cover the whole main episode. Using mail in these examples is only a part of the "buying" and "notifying" episodes. Other expressions similar to "by mail" include "by phone", "by telegraph", "by messenger", etc.

How does the analyzer process BY+INSTR constructions? As opposed to BY+ING constructions where we had to deal with an unlimited variety of actions and where the interpretation of instrumentality was sometimes very fuzzy, handling of BY+INSTR constructions is relatively straightforward. Underlying such words as "bus", "mail", "phone", etc. are instrumental scripts. These scripts specify the activities to which they are instrumental. For example, the $BUS script is instrumental to the physical transfer of people, the $MAIL script is instrumental to the MTRANS of messages and the PTRANS of physical objects. First, the analyzer checks if it is dealing with an instrumental script. Then, if this script can be instrumental to the main act of the sentence, the analyzer expands the act to include the instrument. Consider the following example:

(7.37) JOHN WENT TO NEW YORK BY CAR.

When the parser reaches the preposition "by" in the above sentence, it invokes the prepositional phrase expert which activates the requests stored under this preposition and attempts to collect the noun group that follows. The noun group processor reads the word "car" which can be interpreted either as a vehicle or as the car-driving script. The absence of the article indicates the second. (In languages without articles different means can be used to indicate this distinction. For example, Russian uses the instrumental case to indicate instrumental use of an object.) Thus the noun group processor returns the concept:

(7.38) ($DRIVE VEHICLE (CAR))

This triggers the BY+INSTR request which tries to interpret the above concept as an instrument of the main conceptualization. Since the INSTR slot of the PTRANS suggests a transportation script as a possible filler and $DRIVE is such a script, the BY+INSTR request succeeds.

Not every transportation script should succeed in filling the INSTR slot of a PTRANS. Suppose sentence (7.37) was used while John was in London. The BY+INSTR request should be able to check whether a particular transportation script could be used as an instrument to a particular PTRANS in a given context. This, however, requires a great deal of knowledge about geography and transportation and the ability to represent this knowledge in a form usable by the parser, something that we still do not know how to do. Here again, our goal is not to solve every problem, but to point out the problems that have to be solved.

7.3. Physical objects as instruments.

Although I was trying to limit this excursion through various topics in parsing to the constructions involving the preposition "by", I cannot leave the topic of instrumentality without saying a few words about the third way to introduce instrumentality in English, the "with" prepositional phrases. The "with" prepositional phrases usually introduce a physical object which is used in the main action. We already discussed this situation briefly in Chapter 4. Let us examine to the example used there in greater detail:

(7.39) JOHN BROKE THE WINDOW WITH A HAMMER.

When the analyzer reaches the word WITH, it activates all the experts stored in its definition. One of these experts tries to interpret the next noun group as an instrument for the main action of the sentence which, in this case, is:

```
(7.40) (CAUSE CONCEPT1 (DO ACTOR (JOHN))
              CONCEPT2 (STATE-CHANGE STATE-NAME (PHYSICAL)
                                     OBJECT     (WINDOW)
                                     VAL-FROM   (+10)
                                     VAL-TO     (-10)))
```

The instrumental expert checks the function of the hammer and finds that is to PROPEL things. This conceptualization is substituted for the dummy act DO in (7.40) and the resulting conceptualization is sent to memory to check whether propelling the hammer can indeed cause the physical state change of the window. This approach, however, will not work for the following sentence:

(7.41) JOHN BROKE THE WINDOW WITH A CHAIR.

The function of chairs is for people to sit on. The memory check would return a negative answer: sitting on a chair cannot break the window. Thus, we have to examine our knowledge about breaking things more carefully. There are two standard ways to break a physical object: One is to hit (PROPEL) it with a harder object. The other is to hit a hard object with the first one. When the instrument is specified by the "with" phrase, the second way is ruled out. Hence, the chair is substituted for the hard object in the PROPEL conceptualization. The memory is consulted whether PROPELling a chair can cause the disintegration of a window. This time, the answer is positive.

The reader might ask why we did not apply the second approach in the first place, without checking the function of the physical object. The answer to this objection is that it might lead to incorrect inferences. Consider the following sentence:

(7.42) JOHN GOT THE MONEY WITH A GUN.

If we consult only the ways a person can achieve possession of some object without looking at the function of the gun, we might find that the easiest way to get the money would be to sell the gun. Thus, we would make the inference that John sold his gun, which is not unreasonable. Unfortunately, most people would understand (7.42) differently. They would conclude that John got the money by threatening (THREAT PLAN-BOX) or shooting (OVERPOWER PLAN-BOX) someone, which are the functions of a gun. This means that the standard functions of an object must be considered first.

7.4. Comparison with Russian.

It is interesting to compare instrumental constructions in English with Russian ones. Conceptually, the role of the instrumental constructions in Russian is exactly the same as their role in English, namely to specify the internal structure of the episode underlying the main action of the sentence. However, the tools provided by the Russian language for doing this are, in some sense, much less "precise" than the English ones. Most of the BY+INSTR and BY+PHYSOBJ constructions are expressed in Russian using the appropriate noun in the instrumental case. For example, the instrumental phrases in the following sentences:

(7.43) JOHN WENT TO NEW YORK <u>BY</u> <u>CAR</u>.

and

(7.44) JOHN BROKE THE GATE WITH A CAR.

would be often expressed in Russian with the same word MASHINOI without any prepositions or articles (Russian does not have articles). However, this does not cause any confusion. The instrumental case ending OI indicates that there can be an instrumental connection and the rest is determined on the basis of the speakers' world knowledge.

Russian does not have anything analogous to BY+ING constructions. Sentence (7.45):

(7.45) JOHN ENTERED THE HOUSE BY BREAKING THE WINDOW.

would normally be translated into Russian as:

(7.46) JOHN VOSHEL V DOM RAZBIV OKNO.

which, strictly speaking, does not mean the same thing. What it literally means is "John entered the house having broken the window", which may or may not be interpreted in the instrumental sense, depending on the context. Thus, when it is not clear whether John actually used the window to enter the house, in order to disambiguate sentence (7.46), a Russian speaker has to analyze John's intentions. A good Russian-English interpreter will have to disambiguate (7.46) in order to produce a good English translation, which means that he would have to understand this sentence and make use of his world knowledge and his knowledge about people's goals and plans.

Different means used in Russian to express instrumentality does not mean that Russians understand instrumentality differently than the English speakers. The underlying analysis of instrumental constructions in terms of goals and plans implied by the main action is the same as in English.

7.5. Summary

We have seen that English has three ways to express instrumentality: BY+ING, BY+INSTR, and WITH+PHYSOBJ constructions. As a rule, BY+ING constructions introduce individual actions which specify one of the PLAN-BOXES of the intended goal of the main action, BY+INSTR constructions refer to instrumental scripts, and WITH+PHYSOBJ prepositional phrases introduce physical objects used in the main action. Parsing of instrumental constructions requires a detailed theory of the internal structure of episodes in memory, a more complete and detailed representational mechanism for representing people's goals and plans, and a great deal of world knowledge about geography, communications, etc. In the case of instrumental constructions, as always, a difficulty in parsing is a manifestation of insufficient representational tools and our poor understanding of the memory processes.

CHAPTER 8

MORE ABOUT "BY"

8.0. Introduction.

In this chapter we will describe four constructions involving the preposition "by". These constructions are classified by the concepts underlying the noun group which follows the preposition "by":

BY+PERSON - "John read a novel by Tolstoi", "John walked by Bill".

BY+ORGANIZATION - "John read a report by the Computer Science Department".

BY+CONNECTOR - "John walked home by the old country road", "John entered the house by the back door". CONNECTORs are defined by (Lehnert 1978) as objects which join two regions or spatial locations with respect to a primitive act. The act which is relevant to our discussion is the PTRANS.

BY+PHYSOBJ - "John sat by the tree", "John entered the house by the lake".

Many of the above constructions can be used differently in a sentence. For example, the concept PERSON following the preposition "by" can either specify the author of the MENTAL OBJECT introduced by the last noun group (as in "John read a novel by Tolstoi") or it can give additional information about the path of a moving object (as in "John walked by Bill"). In this chapter we will describe the above constructions, the roles they can play in a sentence, and give some rules for disambiguation between these roles.

Before we proceed further, I would like to introduce some terminology. We will refer to the noun group which follows the preposition "by" as the subject of the construction. The noun group immediately preceding the preposition "by" will be referred to as the last noun group.

8.1. BY+PERSON and BY+ORGANIZATION constructions.

The BY+PERSON and BY+ORGANIZATION constructions can be used in English in four different ways:

PASSIVE+BY - as in "The lecture was given by John", "The book was published by Yale University".

AUTHORSHIP - as in "John read a book by Tolstoi", "John read a report by IBM".

PROXIMITY - as in "John sat by Mary". This construction specifies an object's location to be in the proximity of the subject of the "by" expression.

VIA-PROX - as in "John walked by Bill". This construction specifies that in his PTRANS John passed via a location in the proximity of the subject of the "by" expression.

We already discussed the passive construction in Chapter 6. So it does not need further introduction here.

AUTHORSHIP constructions describe the authorship of an object. For example, we represent sentence:

(8.1) JOHN TOOK A BOOK BY TOLSTOI.

as

```
(8.2) (PTRANS ACTOR  (*JOHN*)
              OBJECT (*BOOK* AUTHOR (*TOLSTOI*))
              FROM   (NIL)
              TO     (*JOHN*))
```

We realize that AUTHORSHIP is not a primitive conceptual category. It can be decomposed according to the ways the object of authorship can be created. When we talk about the authorship of a novel, we mean the person who wrote it. When we talk about the authorship of a song, we might mean one of three things: the person who wrote the words, the person who wrote the music, or the person who made it known by performing it. The authorship of a physical object means either its design or its physical creation. In Section 8.1.2 we will show that this distinction is important in determining whether a person can indeed be the author of an object. However, at the present level of representation we lump all these categories together.

When we say:

(8.3) JOHN SAT BY BILL.

we mean that John had a certain body position and was located near Bill. We represent it as:

```
(8.4)  (STATE STATE-NAME (BODY-POSITION)
              TYPE       (SITTING)
              OBJECT     (*JOHN*)
              SUPPORT    (NIL)
              PLACE      (PROX OBJECT (*BILL*)))
```

We realize that the expression (PROX OBJECT X) is a very poor way of representing locations in the proximity of the object X. It does not make explicit the meaning of "proximity" and has to be interpreted in every particular situation. For example, "proximity between people at a party" means a distance within one or two yards; "proximity between airplanes in the air" means a few miles. However, since we lack a general theory of location representation, we use this notation. In Chapter 9 we discuss some ideas toward a more general solution of this problem.

In sentences describing physical transfers, BY+PERSON and BY+ORGANIZATION constructions can be used to specify the path of the PTRANS. For example:

(8.5) JOHN WALKED BY BILL.

The phrase BY BILL in the above sentence specifies that John's path went in the proximity of Bill. This can be represented as:

```
(8.6)  (PTRANS ACTOR  (*JOHN*)
               OBJECT (*JOHN*)
               FROM   X
               VIA    (PROX OBJECT (*BILL*)
               TO     Y)
```

Thus, the BY+PERSON and BY+ORGANIZATION constructions can play four different roles in a sentence: PASSIVE+BY, AUTHORSHIP, PROXIMITY, and VIA-PROX. The next question is how to determine which role the construction is playing in a particular sentence. For example, the following sentence is four ways ambiguous:

(8.7) THE JEWEL WAS CARRIED TO THE CASE BY CHRISTIAN DIOR.

The following are four different paraphrases of this sentence:

1) Christian Dior carried the jewel to the case.

2) Someone carried the jewel to the case designed by Christian Dior.

3) Someone carried the jewel to the case located near Christian Dior.

4) Someone carried the jewel to the case and on his way passed near Christian Dior.

Our present parsing framework does not have a facility for comparison between different readings of the same expression. In fact, in the design of this framework we consciously tried to avoid the necessity for such a comparison. Instead, we tried to organize parsing hierarchically so that only one hypothesis is tested at any given time. The lower-level hypotheses are tested only when the higher-level ones fail. The assumption underlying this scheme is that in a realistic context one and only one alternative is valid. Let us now examine the hierarchy of rules used in the disambiguation of the BY+PERSON and BY+ORGANIZATION expressions.

8.1.1. PASSIVE+BY constructions

Among the four roles that the BY+PERSON and BY+ORGANIZATION expressions can play in a sentence, the PASSIVE+BY construction stands aside from the others. What distinguishes this construction from other constructions is the fact that it is expected from the processing of the previous parts of the sentence. As described in Chapter 6, in the case of passive sentences we have an expectation for a "by" prepositional phrase which would build a concept filling a slot (usually ACTOR) in the main conceptualization of the sentence. We call this expectation the PASSIVE+BY expectation. The PASSIVE+BY expectation is set up when passive is recognized and is kept on the OPTIONAL line of requests. In the case of other constructions: AUTHORSHIP, PROXIMITY, and VIA-PROX, the processing of the initial part of the sentence does not give us such an "advanced warning". These constructions are not predicted and have to be handled by the appropriate experts from the EXPERT line of requests. Thus, when the parser encounters a BY+PERSON or a BY+ORGANIZATION expression in a passive sentence it first tries to handle it as a PASSIVE+BY construction. It tries the experts from the EXPERT line only if this interpretation fails. (Recall that the OPTIONAL line is higher than the EXPERT line in the hierarchy of requests.) For example, in the following sentence:

(8.8) JOHN WAS GIVEN A BOOK BY HIS PROFESSOR.

the parser interprets the concept underlying the phrase HIS PROFESSOR to be the ACTOR of the ATRANS underlying the main clause. The possibility that the book was written by John's professor would not even be considered.

How do we determine whether a PASSIVE+BY construction is a plausible interpretation of a BY+PERSON or a BY+ORGANIZATION expression? First, the concept built by this expression has to satisfy the PASSIVE+BY expectation. For example, the filler of the ACTOR slot of the ATRANS underlying sentence (8.8) has to be human. This, however, is not sufficient. In order to accept the PASSIVE+BY interpretation, we have to check if it makes sense in the current context. For example, if we knew that John's professor was not present in the context in which sentence (8.8) was used, we would reject the PASSIVE+BY interpretation. Let us examine how this check can be made in a scriptal context by the SAM system (Cullingford 1978). Suppose the parser is reading the following sentence in a

restaurant story:

(8.9) JOHN WAS GIVEN THE MENU BY THE WAITER.

The parser checks if the expectation attached to the ACTOR slot underlying sentence (8.9) is satisfied, fills the slot and sends the conceptualization to the memory module. Thus, the memory module sees the following conceptualization:

```
(8.10) (ATRANS ACTOR  (*WAITER*)
               OBJECT (*MENU*)
               FROM   (*WAITER*)
               TO     (*JOHN*))
```

This conceptualization is matched against the appropriate conceptualizations in the $RESTAURANT script. The match succeeds, the memory module does not complain, and the parse is accepted. If, on the other hand, instead of (8.9) we had:

(8.11) JOHN WAS GIVEN THE MENU BY THE COOK.

THE COOK would satisfy the expectation attached to the ACTOR slot and the parser would send the following conceptualization to the memory module:

```
(8.12) (ATRANS ACTOR  (*COOK*)
               OBJECT (*MENU*)
               FROM   (*COOK*)
               TO     (*JOHN*))
```

The memory module would not be able to match the above conceptualization with any pattern in the script because of the role conflict between the cook and the person who normally gives the menu to the clients. The memory module would complain back to the parser thus forcing it to consider other possibilities (which in this case will also fail).

Scripts are the only type of contexts in which the parser was extensively tested. However, in many cases PASSIVE+BY constructions are used in non-scriptal contexts. Consider the following story:

(8.13) JOHN WAS A STUDENT IN LINGUISTICS AT THE
UNIVERSITY OF ALASKA. HE TOOK A COURSE
ON TRANSFORMATIONAL GRAMMAR. JOHN WAS
TOLD TO READ A BOOK BY CHOMSKY.

An informed reader would consider it highly unlikely that Chomsky personally told John to buy a book in the above story. This reader would know that Chomsky is a professor at MIT and would not expect him to be at the University of Alaska telling a random student to read some unspecified book. This knowledge, however, is not scriptal. Hence, the parser which relies on SAM for its contextual

and the bulk of its world knowledge will not be able to handle this story.

8.1.2. The AUTHORSHIP expert.

In the case of an active sentence or when the PASSIVE+BY interpretation of a BY+PERSON/BY+ORGANIZATION expression fails, the parser "drops" to the EXPERT line of requests which have to decide which of the three remaining roles (AUTHORSHIP, PROXIMITY, and VIA-PROX) this expression can play in the sentence. This is done by two experts: the AUTHORSHIP expert and the BY-LOCATION expert. In this section we will examine the order in which these experts are applied to the BY+PERSON/BY+ORGANIZATION expressions and discuss the AUTHORSHIP expert in greater detail. The BY-LOCATION expert is described in the next section. Consider the following two examples:

(8.14) MARY ATE HER LUNCH BY THE COMPUTER SCIENCE DEPARTMENT.

and

(8.15) MARY WAS CARRYING A NEW EXPERIMENTAL COMPUTER BY THE COMPUTER SCIENCE DEPARTMENT.

The first sentence means that Mary ate near the Department, rather than that her lunch was made by the Computer Science Department. This is because we know that Computer Science Departments do not make lunches. However, even the second sentence would be usually interpreted as Mary's location, although the computer science department was a possible creator of her computer. Compare this with the following sentence:

(8.16) MARY READ A REPORT BY THE COMPUTER SCIENCE DEPARTMENT.

This sentence means that the report was created by the Computer Science Department, although we can think of the context which would force the first interpretation. This will be the case when the context clearly states that the report that Mary was reading was not produced by the Computer Science Department. The above examples suggest different biases for mental and physical objects. The following rule takes this into account:

If the preposition BY follows a physical or mental object and is followed by a specification of a person or an organization then: if the object is mental then first try to interpret the expression as an AUTHORSHIP construction; if the object is physical then try to interpret the expression as a location specifier first.

For example, in the following sentence:

(8.17) MARY BOUGHT A BOOK BY CHOMSKY.

the book is primarily a mental object and, hence, the phrase "by Chomsky" indicates the authorship of the book. When we hear:

(8.18) MARY WAS LYING WITH A BOOK BY CHOMSKY.

we do not consider the possibility that the phrase "by Chomsky" indicates Mary's location with respect to the professor. If, on the other hand, we had:

(8.19) MARY WAS LYING WITH A BOOK BY HER HUSBAND.

we would probably notice the ambiguity and then, if we do not know whether Mary's husband writes books or not, interpret the BY phrase as Mary's location since we know from our world knowledge that writing is a rather exclusive profession and a random man is not likely to be one.

If, on the other hand, we had:

(8.20) MARY WAS LYING ON THE BED BY HER HUSBAND.

we would not normally consider the possibility that the bed was made by her husband, although such a possibility certainly exists. The hypothesis that in the case of physical objects we first check the locational interpretation can explain why the following sentence usually provokes smiles:

(8.21) ROCKEFELLER WAS LYING ON THE BED BY DALI.

The bed is a physical object. Therefore, we try to interpret the phrase BY DALI as a location specifier. Only when this interpretation is rejected as implausible, do we check and indeed might remember that Rockefeller had a bed designed by Salvador Dali.

Suppose now that we have a hypothesis that a particular construction is an AUTHORSHIP expression. How can we verify it? There are two problems that the analyzer has to solve at this point. First, it has to determine the meaning of the AUTHORSHIP expression (we have seen that authorship can mean different things), and,

second, it has to verify that this type of authorship makes sense in the current context. We do not have a general solution to these problems. What follows is an attempt to identify the kinds of information the parser would need in order to decide whether a BY+PERSON/BY+ORGANIZATION expression can be interpreted as an AUTHORSHIP construction.

The first problem of how to determine the meaning of an AUTHORSHIP expression has a relatively straightforward solution. Every object determines the primary meaning of its authorship. For example, books are written, films are directed, clothing is designed, etc. This seems to be a little too specific, however, we cannot use anything more general. For example, the reference to the principal participants in the creation of the object will not do. A film is created by a script writer, a director, cameramen, actors, and a producer. However, when we say "a film by Fellini" we mean the director and not the script writer or the producer. Thus,

|---|
| In order to parse AUTHORSHIP expressions we |
| need to consult knowledge about physical and |
| mental objects. This knowledge normally includes |
the primary meaning of objects' authorship.

Some mental objects have several standard meanings for their authorships. For example:

(8.23) A SONG BY JOHN DOE

can have three interpretations: John Doe wrote the music; John Doe wrote the lyrics; and John Doe performed it. The disambiguation of this expression requires additional information about John Doe. If we knew that John Doe was a poet, for example, then we would conclude that he was the author of the lyrics.

Each type of authorship places certain requirements on the potential author. An author of music, for example, has to have certain musical skills. An author of a letter has to know how to write. Violation of these requirements leads to the rejection of the particular authorship interpretation. For example:

(8.24) THE DESTRUCTION OF A BOOK BY THE BABY.

In order to be the author of a book one at least has to know how to write. On the other hand, we know that babies do not have this skill. Therefore, the authorship interpretation of (9.33) is rejected. Thus,

In order to parse AUTHORSHIP expressions we need to consult knowledge about the skills of the potential author of the object. These skills have to match one of the meanings of the objects' authorship.

In order to handle AUTHORSHIP expressions, in addition to the information about the ways the object can be created we have to have two other kinds of information: (1) information about the subject's occupation and (2) information about the current context. We need the first kind of information because the authorship of the object has to correspond to something in the subject's occupation. Otherwise we might prefer another interpretation. For example, when we hear:

(8.25) MARY WAS GIVEN A DRESS BY HER PROFESSOR.

we understand it to mean that Mary's professor gave her a dress, since making dresses is not a part of the PROFESSOR ROLE-THEME. If, on the other hand, we hear:

(8.26) MARY WAS GIVEN A DRESS BY CHRISTIAN DIOR.

we assume that Mary's new dress was designed by the famous designer, since designing dresses is a part of the DESIGNER ROLE-THEME. Thus,

In order to parse AUTHORSHIP expressions we need to consult knowledge about the occupation of the subject of the expression. If making the objects described by the object of the expression is part of his occupation, the expression should be interpreted as an AUTHORSHIP specification.

In addition to our general knowledge about the subject's ROLE-THEME, we might know a few specific facts about him which might affect the interpretation of an AUTHORSHIP expression. For example, in addition to the fact that Chomsky is a professor which means that he teaches classes, advises students, writes papers, goes to conferences, etc., we know a few specific facts about Chomsky personally, for example, that he invents constructions called transformations and writes books about them. For example, this information is used to reject the authorship interpretation in the following sentence:

(8.27) MARY WAS TOLD TO READ A BOOK ON ASTRONOMY
BY CHOMSKY.

since we know that Chomsky is a professor of linguistics and has

never written a book on astronomy.

Finally, the analyzer has to be able to check whether the interpretation makes sense in the current context, which, in this case means the verification that the concrete object (not the conceptual type) could be produced by the concrete subject of the expression in the current context. In the following sentence, for example,

(8.28) MARY READ HER REPORT BY THE COMPUTER SCIENCE DEPARTMENT.

we reject the authorship interpretation although reports can be produced and are normally produced by computer science departments. However, in the particular context created by sentence (8.28), we already knew that the report was produced by Mary and, therefore, could not be produced by the Computer Science Department.

Thus, the correct handling of AUTHORSHIP expressions requires knowledge of the ways various objects can be created, knowledge of ROLE-THEMES, and knowledge of the specific context of the situation. At the moment, we cannot give any more specific rules of how to access this knowledge, since this would require much more specific representational mechanisms for these knowledge structures than the ones that are presently available.

8.1.3. The BY-LOCATION expert.

Let us first summarize the actions of the parser described in the two previous sections. We will do it in the form of an informal algorithm. When the parser encounters a BY+PERSON/BY+ORGANIZATION expression it proceeds as follows:

1) If the BY+PERSON/BY+ORGANIZATION construction follows an object and the main function of the PERSON/ORGANIZATION is creating this type of objects, then try the AUTHORSHIP interpretation, otherwise go to step 2. (Thus, when the parser reads: "John was driven to the house by the famous architectural firm" it first tries to interpret the "by" phrase as the AUTHORSHIP construction and tries other interpretations only if the first interpretation fails.)

2) If the sentence is passive then go to step 3, otherwise go tc step 4.

3) Try the PASSIVE+BY interpretation. If the interpretation succeeds exit, otherwise go to step 4.

4) If the BY+PERSON/BY+ORGANIZATION construction follows a mental object then try the AUTHORSHIP interpretation. If the interpretation succeeds exit, otherwise go to step 5.

5) Try PROXIMITY and VIA-PROX interpretations. If one of them succeeds exit, otherwise go to step 6.

6) If the BY+PERSON/BY+ORGANIZATION construction follows a physical object then try the AUTHORSHIP interpretation.

In this section we concentrate on step 4 of the above procedure, namely the disambiguation between the PROXIMITY and VIA-PROX constructions. In these constructions we consider people and organizations as simple physical objects. Thus the following discussion is also relevant to the BY+PHYSOBJ constructions. The disambiguation between PROXIMITY and VIA-PROX constructions is done by the BY-LOCATION expert.

By definition the VIA-PROX construction can be used only in PTRANS conceptualizations. Thus, when the concept underlying the main clause of the sentence is not a PTRANS, the parser interprets BY+PERSON, BY+ORGANIZATION, and BY+PHYSOBJ expressions as PROXIMITY constructions. Let us examine this type of sentences in greater detail.

8.1.3.1. Non-PTRANS Sentences.

Consider the following two examples:

(8.29) JOHN BOUGHT A HOTDOG BY THE LIBRARY.

and

(8.30) JOHN BOUGHT A HOUSE BY THE LIBRARY.

In the first sentence, the phrase BY THE LIBRARY specifies the location of John while buying the hotdog. In the second sentence this phrase specifies the house that John bought by giving its location. How do we disambiguate between these two interpretations? The difference between the above examples is that houses have fixed locations while hotdogs do not. Thus, we propose the following rule for PROXIMITY constructions:

```
|-----------------------------------------------|
| If a BY+ORGANIZATION  or a BY+PHYSOBJ expression |
| follows  an object  which normally  has  a fixed |
| location it specifies the location of the object,|
| otherwise it specifies the location of the act.  |
|-----------------------------------------------|
```

The assumption underlying the above rule is that people do not normally characterize movable objects by their locations. It is really hard to force such a characterization. Even in the following story:

(8.31) THERE WERE TWO ICE-BOXES WITH BEER. ONE
BY THE TREE AND THE OTHER BY THE BUSH.
JOHN DRUNK ALL THE BEER BY THE BUSH.

we are inclined to assume that John was near the bush while he was drinking all this beer.

If the BY+ORGANIZATION/PHYSOBJ/PERSON expression does not follow a noun group or if the last noun group is an object without a fixed location in the current context then the PROXIMITY construction built by the "by" phrase specifies the location of the act or state underlying the sentence (we are still considering non-PTRANS conceptualizations). For example:

(8.32) ON THE BUS JOHN SAT BY THE OLD LADY.

Application of the above heuristic often depends on the current context. Objects that do not have fixed locations in one setting can be considered fixed in another. Compare, for example, the following sentences:

(8.33) JOHN BOUGHT THE CHRISTMAS TREE
BY THE CITY LIBRARY.

and

(8.34) JOHN DECORATED THE CHRISTMAS TREE
BY THE CITY LIBRARY.

Christmas trees are sold in many different places. Thus, in a Christmas trees sale context, the trees are movable objects without fixed locations. The phrase BY THE CITY LIBRARY in the first sentence specifies the location of the transaction (the one near the library). On the other hand, trees that are being decorated normally have fixed locations, therefore, the same phrase in the second sentence specifies the location of the tree.

8.1.3.2. PTRANS Sentences.

Let us recapitulate the main steps of the process that we are describing. We assume that the parser has encountered a BY+PERSON or a BY+ORGANIZATION expression. Further, we assume that either the sentence in which this expression was encountered is active or, if it was passive, that the PASSIVE+BY interpretation of the expression has been rejected. We also assume that the AUTHORSHIP interpretation has been rejected as well. Thus, the only alternatives left to the parser are the PROXIMITY and the VIA-PROX constructions. (We remind the reader that VIA-PROX constructions specify the path of PTRANS's.) As we noted above, the VIA-PROX construction can be used only in sentences with the underlying PTRANS conceptualization. In the last section we examined the

PROXIMITY construction as used in sentences whose underlying conceptualization is not a PTRANS. In this section we examine PROXIMITY and VIA-PROX constructions in PTRANS conceptualizations.

When BY+PERSON or BY+ORGANIZATION expressions do not follow a noun group the parser interprets them as VIA-PROX constructions. For example:

(8.35) JOHN WALKED BY BILL.

or

(8.36) JOHN RAN BY THE LIBRARY.

The first of the above sentences means that John passed by Bill, the second that John passed by the library. We represent the meaning of the first of the above sentences as:

```
(8.37) (PTRANS ACTOR  (*JOHN*)
               OBJECT (*JOHN*)
               INSTR  ($WALK)
               FROM   X
               VIA    (PROX OBJECT (*BILL*))
               TO     Y)
```

and the meaning of the second as:

```
(8.38) (PTRANS ACTOR  (*JOHN*)
               OBJECT (*JOHN*)
               INSTR  ($RUN)
               FROM   X
               VIA    (PROX OBJECT (*LIBRARY*))
               TO     Y)
```

Let us now consider the case when a BY+PERSON or BY+ORGANIZATION expression is used following another noun group. Here we have to distinguish two cases. One is when the concept underlying the last noun group fills the TO or FROM slots of the PTRANS and the case when the last noun group before the "by" phrase fills the OBJECT slot. In the first case we normally have a PROXIMITY construction, in the second a VIA-PROX. For example:

(8.39) JOHN WALKED TO THE ROCK BY THE LIBRARY.

The phrase BY THE LIBRARY in the above sentence specifies the location of the rock. However, the same phrase in the following sentence:

(8.40) JOHN CARRIED THE ROCK BY THE LIBRARY.

specifies the path of John's PTRANS. Thus,

```
|-------------------------------------------------|
| If a BY+ORGANIZATION  or a BY+PHYSOBJ expression |
| follows  an object  which fills  the TO  or FROM |
| slots  of  the PTRANS conceptualization  then it |
| specifies the location of the object  (PROXIMITY |
| construction);   if  it follows  an object which |
| fills the OBJECT slot then it specifies the path |
| of the PTRANS (VIA-PROX).                        |
|-------------------------------------------------|
```

This rule is based on the fact that objects that are being moved are not normally characterized by their location.

The BY+PERSON construction is rarely used in PROXIMITY or VIA-PROX constructions in PTRANS conceptualizations. The following sentences, for example, sound somewhat odd:

(8.41) JOHN WALKED TO THE ROCK BY THE OLD LADY.

(8.42) JOHN CARRIED THE ROCK BY THE OLD LADY.

People do not have fixed locations and, therefore, are not normally used to specify other objects' locations. However, sentences (8.41) and (8.42) are understandable and our heuristic will work on them.

The following diagram summarizes the work of the BY-LOCATION expert:

```
                           |---------|
                           | PTRANS? |
                           |---------|
                  Yes         |     |        No
            |-------------|     |-----------|
            V                               V
 |----------------------|     |-----------------------|
 | LASTNG is a PHYSOBJ  |     | LASTNG is a PHYSOBJ    |
 | filling:             |     | with a fixed location?|
 |----------------------|     |-----------------------|
   |                  |         |                  |
   | FROM or TO       | OBJECT  | Yes              | No
   | slots            | slot    |                  |
   V                  V         V                  V
PROXIMITY         VIA-PROX   PROXIMITY          PROXIMITY
(LASTNG)                       (ACT)             (LASTNG)
```

Figure 8.1: BY-LOCATION expert

8.1.4. Summary.

Let us summarize the processing of the BY+PERSON and BY+ORGANIZATION expressions. When the parser reads the preposition "by" in the structure-driven mode it first checks the requests on the ACTIVE and OPTIONAL lines. For example, if the sentence is passive one of the requests on the OPTIONAL LINE would be expecting a "by" prepositional phrase (see Chapter 6). If all the requests on these lines fail, the parser "drops" to the EXPERT line of requests and the prepositional phrase expert (see Chapter 4) is activated. This expert collects the noun group following the preposition (if it has not already been collected by the PASSIVE+BY request) and activates requests in the definition of the preposition "by". At present, we have six such requests: BY+SCRIPT (described in Chapter 7), BY+PERSON, BY+ORGANIZATION, BY+CONNECTOR, BY+PHYSOBJ, and BY+TIME (described in Chapter 9). Thus, if the underlying conceptualization of the noun group following the preposition "by" is a PERSON, the BY+PERSON request is triggered, if it is an ORGANIZATION, BY+ORGANIZATION request is triggered. (The last four requests will be described shortly.) The BY+PERSON and BY+ORGANIZATION requests check if there was a noun group preceding the preposition "by". If there was none, they try the BY-LOCATION expert. If the concept underlying the last noun group was a mental object, they try the AUTHORSHIP expert first, and then try the BY-LOCATION expert if the first one fails. If the concept underlying the last noun group was a physical object, the BY-LOCATION expert is tried first and the AUTHORSHIP expert second. The actions of the BY+PERSON and BY+ORGANIZATION requests can be represented schematically as follows:

```
                   |-----------|
                   | LASTNG =  |
                   |-----------|
    Mental Object    |    |    |   Physical Object
        |------------|    |    |-----------|
        |                 | nil            |
        V                 V                V
  |------------|  |-------------|  |-------------|
  | AUTHORSHIP |  | BY-LOCATION |  | BY-LOCATION |
  | expert     |  | expert      |  | expert      |
  |------------|  |-------------|  |-------------|
        |                                  |
        | fail                             | fail
        V                                  V
  |-------------|                  |------------|
  | BY-LOCATION |                  | AUTHORSHIP |
  | expert      |                  | expert     |
  |-------------|                  |------------|
```

Figure 8.2: BY+PERSON and BY+ORGANIZATION requests

8.2. BY+CONNECTOR and BY+PHYSOBJ expressions.

In this section we consider the following type of sentences:

(8.43) JOHN ENTERED THE HOUSE BY THE BACK DOOR.

We first assume that the parser encountered the preposition "by" followed by a noun group whose underlying concept is a CONNECTOR. We also assume that none of the ACTIVE or OPTIONAL line requests were satisfied and the parser "dropped" to the EXPERT line at which point the requests stored under the preposition "by" were activated by the prepositional phrase expert and the BY+CONNECTOR request was triggered. At this point the parser can encounter the following possibilities:

1) The concept underlying the "by" phrase is used as a CONNECTOR with respect to the PTRANS underlying the sentence (for example, sentence (8.43)). Doors are connectors with respect to PTRANS's in and out of enclosed areas. Entering a house is such a PTRANS. We represent the meaning of sentence (8.43) as:

```
(8.44) (PTRANS ACTOR     (*JOHN*)
               OBJECT    (*JOHN*)
               FROM      (OUTSIDE OBJECT (*HOUSE*))
               CONNECTOR (*BACK-DOOR*)
               TO        (INSIDE OBJECT (*HOUSE*)))
```

2) The second possibility is when the connector is not used as such but describes a location in a PROXIMITY or VIA-PROX construction. For example:

(8.45) JOHN SAT BY THE OLD COUNTRY ROAD.

(8.46) JOHN BOUGHT A HOUSE BY THE BRIDGE.

(8.47) JOHN WALKED BY THE WINDOW.

The first possibility is handled by the CONNECTOR expert. The second possibility can be handled by the BY-LOCATION expert described in section 8.1.3.

When the parser encounters a BY+CONNECTOR expression, the BY+CONNECTOR request gets triggered. This request first calls the CONNECTOR expert to check if the expression can be interpreted as a connector with respect to the act underlying the sentence. If this attempt fails, the BY+CONNECTOR request decides that the expression was used as a location specifier and calls the BY-LOCATION expert. In the case of "by" prepositional phrases whose underlying conceptualizations are physical objects which are not connectors, the BY+PHYSOBJ request gets triggered which calls the BY+LOCATION expert directly, without trying the CONNECTOR interpretation first. For example, houses are not normally used as connectors, therefore the handling of the "by" phrase in the following sentence:

(8.48) JOHN PLANTED A TREE BY THE HOUSE.

is passed to the BY-LOCATION expert which interprets it as a PROXIMITY construction specifying the location of the tree which John planted. The following diagram represents the actions of the BY+CONNECTOR and BY+PHYSOBJ requests schematically:

```
      BY+CONNECTOR                      BY+PHYSOBJ
        request                          request

     |-----------|                   |-------------|
     | CONNECTOR |                   | BY-LOCATION |
     | expert    |                   | expert      |
     |-----------|                   |-------------|
           |
           | fail
           V
    |-------------|
    | BY-LOCATION |
    | expert      |
    |-------------|
```

Figure 8.3: BY+CONNECTOR and BY+PHYSOBJ requests

Let us now examine the CONNECTOR expert in greater detail. This expert has to examine the conceptualization built by the sentence and the connector built by the "by" phrase and determine whether the connector can be used as such with respect to the act in the main conceptualization in the current context. For example, in the following sentence:

(8.49) JOHN DROVE HOME BY I-95.

the conceptualization underlying JOHN DROVE HOME is:

```
(8.50) (PTRANS ACTOR     (*JOHN*)
               OBJECT    (*JOHN*)
               FROM      X
               TO        (PROX OBJECT (*HOUSE*))
               INSTR     ($DRIVE))
```

The concept underlying I-95 is a LINK which can be somewhat simplistically represented as:

```
(8.51) (#LINK TYPE          (*HIGHWAY*)
       ^      NAME          (I-95)
       |      CONNECTOR-
       |          ENABLE    (PTRANS ACTOR W OBJECT W
       |                            INSTR ($DRIVE)
       |                            FROM  Y TO     Z
       |                            CONNECTOR . ))
       |                                       |
       |---------------------------------------|
```

The CONNECTOR expert has to match the CONNECTOR-ENABLE conceptualization of the LINK I-95 with conceptualization (8.50). In our case the match succeeds and the CONNECTOR expert produces:

```
(8.52) (PTRANS ACTOR      (*JOHN*)
               OBJECT     (*JOHN*)
               FROM       X
               CONNECTOR  (#LINK TYPE (*HIGHWAY*)
                                 NAME (I-95))
               TO         (PROX OBJECT (*HOUSE*))
               INSTR      ($DRIVE))
```

On the other hand, in the following sentence:

(8.53) JOHN DROVE HOME BY THE RIVER.

the CONNECTOR expert does not succeed. Rivers are connectors only with respect to boating or swimming. The river in the above sentence will be represented as:

```
(8.54) (#LINK TYPE          (*RIVER*)
       ^      CONNECTOR-
       |          ENABLE    (PTRANS ACTOR W OBJECT W
       |                            INSTR ($BOAT, $SWIM)
       |                            FROM  Y TO     Z
       |                            CONNECTOR . ))
       |                                       |
       |---------------------------------------|
```

The PTRANS in (8.54) does not match the PTRANS underlying JOHN DROVE HOME (conceptualization 8.50) because the filler of the INSTR slot in the latter ($DRIVE) does not match any of the possible fillers of the same slot in the former. Thus, the CONNECTOR expert fails and the BY-LOCATION expert is considered.

The above examples were relatively straightforward. The CONNECTOR expert used general knowledge about the types of PTRANS's that roads and rivers support in order to accept or reject the connector interpretation of the above sentences. However, often this simple information is not sufficient, and the CONNECTOR expert needs access to much more detailed knowledge about the physical setting of the current episode. Consider the following examples:

(8.55) JOHN ENTERED THE HOUSE BY THE BRIDGE.

and

(8.56) JOHN ENTERED THE CASTLE BY THE BRIDGE.

The first sentence would be usually understood as "John entered the house which was located near the bridge". The second sentence is truly ambiguous out of context. The castle could be located near the bridge, or the bridge could serve as an entrance to the castle. In order to handle the above sentence the parser must have knowledge about bridges. For example, it must know that bridges are connectors with respect to PTRANS's from one side of a body of water to the other. The parser also must know that physical transfers from the outside of a house to its inside do not normally require crossing a body of water. Castles, on the other hand, are frequently surrounded by moats which require bridges to cross them.

In a specific context we could have a different picture. For example, the house referred to in sentence (8.55) could be located on an island using a bridge as an entrance, in which case the BY THE BRIDGE phrase would have been interpreted as a CONNECTOR. The castle in sentence (8.56) might either have or not have a moat which would entail different interpretations of this sentence. Thus, in order to make the CONNECTOR expert work we have to be able to represent houses, castles, bodies of water, bridges, etc. within a coherent picture of local geography. At present our representational mechanisms are not sufficiently developed to handle this type of information.

Suppose now that the connector interpretation of a BY+CONNECTOR expression was rejected and the BY-LOCATION expert interpreted it as a VIA-PROX construction. In the case where the connector in question is an extended object we can still have more than one interpretation. For example, consider:

(8.57) JOHN WALKED BY THE RIVER.

The most common interpretation of the above sentence is that John walked along the river. However, we can also interpret this sentence to mean that on his way, John was near the river at some point (maybe even crossing it). There is no information in the above sentence which can help us to choose between these interpretations. The disambiguation of this sentence depends on knowledge of local geography and John's starting and finishing locations.

Thus, in order to interpret and properly attach BY+CONNECTOR and BY+PHYSOBJ expressions, the parser has to have access to knowledge about general properties of things (such as that houses have back doors, castles may have bridges, etc.), the internal structure of actions (such as walking, driving, etc.), and the important properties of objects such as their usual function. For example, doors are connectors between the outside and the inside of a house, boats are means of transportations on water, etc. In

addition to this, the parser has to have knowledge about the physical setting of the current episode which includes local geography, properties of the particular objects used in the episode, and the particular actions of the participants.

8.3. Summary

In this chapter, we discussed the processing of several classes of unpredicted "by" prepositional phrases. We classified these phrases according to the concepts underlying the noun group which follows the preposition "by" and described the parsing experts which handle each of these classes. The following diagram shows these classes and the experts used in their processing:

```
|--------|                 |-------------|
| PERSON |--------|        | Follows an  |
|--------|        |        | object and  | Yes  |------------|
                  |-->     | is a usual  |----->| AUTHORSHIP |
|--------------|  |        | creator of  |      | expert     |
| ORGANIZATION |--|        | this type   |      |------------|
|--------------|           | of objects? |            |
                           |-------------|            | fail
                              |                       |
                              V No                    |
          Yes      |----------|                       |
     |-------------| PASSIVE? |<----------------------|
     |             |----------|
     |                  |                   |-------------|
     |             No   |   |---------| No  | BY+LOCATION | fail
     |                  |-->| Follows |---->| expert      |--|
     V                      | mental  |     |-------------|  |
|--------------|       |->  | object? |--|                   V
| BY+PASSIVE   |-------|    |---------|  | Yes       |------------|
| expectations | fail                    V           | AUTHORSHIP |
|--------------|                 |------------|      | expert     |
                                 | AUTHORSHIP |      |------------|
                                 | expert     |--|
                                 |------------|  |---------| fail
                                                           |
|-----------|   |---------| Yes |--------------|           |
| CONNECTOR |-->| PTRANS? |---->| BY+CONNECTOR |           |
|-----------|   |---------|     | expert       |           |
                    |           |--------------|           |
|---------|         | No               |                   |
| PHYSOBJ |-------->|                  | fail              |
|---------|         |                  V                   |
                    |              |--------------|        |
                    |------------->| BY+LOCATION  |<-------|
                                   | expert       |
                                   |--------------|
```

Figure 8.4: "By" prepositional phrase

CHAPTER 9

BY+TIME EXPRESSIONS

9.0. Introduction.

The preposition "by" followed by a time specifying expression gives the "upper" time boundary of an activity. For example:

(9.1) JOHN FINISHED THE WORK BY TUESDAY.

(9.2) BY NOW HE IS IN LONDON.

(9.3) HE HAS TO BE BACK IN NEW YORK BY A WEEK FROM TODAY.

Although the activity bounded by a BY+TIME construction does not have to be finished exactly at the time specified in it, there is a strong implication that it was or would be finished not long before this time. For example, we would not use (9.1) if John had finished the work on Sunday. Thus,

```
|---------------------------------------------|
| BY+TIME constructions imply that the        |
| actual upper time limit of the activity     |
| specified in the sentence immediately       |
| precedes the time given in the construction.|
|---------------------------------------------|
```

When BY+TIME expressions are used in a sentence with an action underlying the main act, it gives the upper boundary of this action. However, when these expressions are used with a state, the action which caused this state has to be inferred. For example, sentence (9.2) implies that there was a PTRANS which resulted in John's being in London and that this PTRANS was completed immediately before the time of the uttering of this sentence. Consider another example:

(9.4) THE CHAIR WAS RED ON TUESDAY.

The above sentence simply specifies the color of the chair at a some point in time. This can be represented in CD as follows:

```
(9.5)  (STATE STATE-NAME (COLOR)
              OBJECT  (*CHAIR*)
              VALUE   (RED)
              TIME    (TIME TIME-TYPE (DAY) WEEK-DAY (3)))
```

(As a side point, notice that the past time specification on a STATE which describes a normally constant property of an object usually implies that this object is not in this state anymore.) Sentence (9.6):

(9.6) THE CHAIR WAS RED BY TUESDAY.

on the other hand, implies that the chair was painted red on or shortly before Tuesday. Thus, we have the following rule about BY+TIME constructions:

```
|-----------------------------------------|
| BY+TIME constructions used with STATE   |
| conceptualizations give the upper time  |
| boundaries on the actions which caused  |
| these states.                           |
|-----------------------------------------|
```

The above rule, however, does not say anything about how the actions which cause a STATE can be inferred. The answer to the latter problem requires detailed knowledge about states and actions which can achieve them. Just to illustrate the difficulties let us consider the following example:

(9.7) THE APPLES WERE RED BY SEPTEMBER.

The above example shows that simple association of color with painting would not be sufficient. We know that, normally, no one paints apples and that apples become red through the process of ripening.

The upper time boundaries of an event can be used in a great variety of inferences depending on the nature of the event and the current context. For example, if we knew that John had to finish some important work before Wednesday and then hear sentence (9.1) which says that he did it by Tuesday, we would infer that John's mental state was high, that whoever gave him the job was happy, etc. If, on the other hand, the job was supposed to be finished before Sunday and John finished it after the deadline, we would make different inferences about John's and his employer's mental states. However, in order to parse BY+TIME expressions and make all these inferences, we have to be able to represent time boundaries, which means that we have to be able to represent points in time and time intervals. In the rest of this chapter we describe a system for time representation and show how natural language time specifications can be parsed into their representations in this system.

9.1. Representing Time.

Our system for representing time was developed in connection with our work on understanding simple script-based newspaper stories which do not require very elaborate time-keeping systems. The system that we are about to sketch is far from being exhaustive and handles only a few types of time specification. A much more elaborate time-keeping system is now being designed. Thus, what follows should be considered only a first pass designed to locate and identify the problems that need to be solved.

Before we proceed any further I would like to add another disclaimer. Every time-keeping system has to have at least two parts: a representational mechanism and a system of inference rules which can use these representations. Since our primary concern is parsing, I will say almost nothing about the inference rules which are needed to make our representations work.

Two approaches to time representation have been discussed in the literature. One proposes decomposition of time specifications in terms of some absolute time scale, such as the number of milliseconds from some point in time (see, for example, Bruce 1972, Winograd 1972). The advantage of this approach is that every point in time can be represented in one and only one way, i.e., that this representation is canonical. This approach has serious disadvantages, however. The first is the fact that in the real world time specifications are "fuzzy" and it is often impossible to translate them into some point on an absolute time scale. For example, we can talk about an event, say a business meeting, happening on February 19, or in the morning of February 19, or on February 19, at 10am. We can also think about the same event as a point in time or as a time interval, a fact that an absolute time scale representation system would not be able to capture. In addition to its inability to deal with the fuzziness of time specifications, such a system would not be able to handle such time specifications as "a day in September" or "a week in February". These time specifications do not correspond to any particular point on the time scale.

The second approach to time representation (Kahn and Gorry 1977) does not use a single absolute time scale. Instead, it uses basic units such as a day, a month, and a year. It also uses a "fuzz" parameter with every date specification. For example, the month of February 1979 would be represented in this system as:

(9.8) (DATE (1979 2 NIL) (FUZZ (MONTHS 1)))

In our system of time representation we use a similar approach. The general form of our representation of a point in time is:

(9.9) (TIME TYPE <time-type> <<unit> <value>>*)

where the time-type is one of eight common time categories:

<time-type> = HOUR | DAYPART | DAY | WEEKEND | WEEK | MONTH | SEASON | YEAR

We use the following time units and their values:

Time units	Values
HOUR	0 < number <,= 2400
DAYPART	MORNING | AFTERNOON | EVENING | NIGHT
DAY	1 <,= number <,= 31
WEEK-DAY	1 <,= number <,= 7
WEEK	1 <,= number <,= 5
MONTH	1 <,= number <,= 12
SEASON	WINTER | SPRING | SUMMER | FALL
YEAR	number

Figure 9.1: Default values of time-units

For example, "a Sunday in September" can be represented in this system as:

(9.10) (TIME TYPE (DAY) WEEK-DAY (1) MONTH (9))

In addition to the values listed above, each time-unit slot can be filled with a time construction corresponding with a corresponding time-type. This type of construction is used in relative time specifications which will be discussed shortly.

Each time-type determines the units which can be used in its specification and also implicitly specifies the "fuzziness" of the time point. For example, the time-type DAY can be described by the following time-units: DAY, WEEK-DAY, WEEK, MONTH, SEASON, and YEAR. It cannot be specified by HOUR, or DAYPART. The following diagram shows time-units which can specify the above time-types. (The top line shows time-units.)

```
HOUR  DAYPART  DAY  WEEK-DAY  WEEK  MONTH  SEASON  YEAR
 |      |       |             |      |       |       |
 |      |       |             |      |       |      YEAR
 |      |       |             |      |       |       |
 |      |       |             |      |       |-------|
 |      |       |             |      |      SEASON   |
 |      |       |             |      |               |
 |      |       |             |      |----MONTH------|
 |      |       |             |                      |
 |      |       |             |----WEEKEND, WEEK-----|
 |      |       |                                    |
 |      |       |---------DAY------------------------|
 |      |                                            |
 |      |------DAYPART-------------------------------|
 |                                                   |
 |-----HOUR------------------------------------------|
```

Figure 9.2: Time-type specifications

We treat the above time-types as primitives rather than decomposing them every time into some other time units for two reasons: First, such decomposition is very difficult because of the fuzziness of these concepts. The second, and most important reason, is these time-types organize rich sets of inferences about time. For example, on weekends people usually do not work, they have parties, visit relatives, etc. During the daytime people usually work (unless it is a weekend), offices and businesses are opened, it is daylight outside, etc. It would be much harder to make all these inferences starting with some absolute time keeping system which represent points in time as the number of seconds elapsed since some fixed origin. The disadvantage of this approach is that the representations it uses are not canonical. Time-keeping systems which use non-canonical representations must have a system of inference rules which can keep the time data base consistent. The discussion of all these inferences lies outside of the scope of this thesis.

So far we have discussed only one type of time specification, namely absolute time specifications such as "February 20". In addition to absolute time specifications a point in time can be specified relative to some other point in time. We distinguish four types of relative time specifications:

1) A point in time can be specified through some event which occurred during that time. For example: "the hour of the earthquake", or "the week of the accident". This type of time specification presupposes that the referent was either equal in duration or shorter than the time-type it specifies. We call this type of time specification <u>sub-referent</u> because the referent event is a subpart of the point in time being specified.

2) A point in time can be given as a subpart of a larger time interval. For example: "the morning of February 21", "five in the afternoon", or "a week in March". We call this type of time specification <u>super-referent</u> because the referent event includes the point in time being specified.

3) A point in time can be specified in relation to an earlier event. For example: "three days after February 19", or "a week from now". We call this type of time specification <u>post-referent</u> because it specifies an event which happens after the referent event.

4) A point in time can be specified in relation to a later event. For example: "three days before the accident", or "a year ago". We call this type of time specification <u>pre-referent</u> because it specifies an event which happens before the referent event.

The following diagram illustrate the above types of relative time specification:

```
                              event
                            V       V
          -----------------|-------|-----------------
          ^           ^       ^       ^            ^
referent  |           |       | 1     |            |
events    |           |               |            |
          |           |---------------|            |
          3                   2                    4
```

Figure 9.3: Types of time specification:
(1) sub-referent, (2) super-referent
(3) post-referent, (4) pre-referent

We represent the sub-referent type of time specifications using the SUB-REF link. For example, the following expression:

(9.11) THE YEAR OF THE ACCIDENT

can be represented as:

(9.12) (TIME TIME-TYPE (YEAR) SUB-REF ($ACCIDENT))

where $ACCIDENT is a pointer to an accident script. If we knew the exact year of the accident, say 1970, we could represent the above expression directly, without referring to the accident:

(9.13) (TIME TIME-TYPE (YEAR) YEAR (1970))

Parsing of the sub-referent type of time specifications is relatively straightforward. The expressions which give this type of time specifications usually consist of a noun group which specifies the time-type followed by an "of" prepositional phrase or a relative

subclause describing the referent event.

The following expression represents a more difficult case of sub-referent time specifications:

(9.14) THE TIME OF THE CAR ACCIDENT

The above phrase can mean the time of the collision proper (order of seconds), the time immediately after until some rescuers or authorities arrive (order tens of minutes), the time needed to liquidate the immediate consequences of the accident such as removal of the injured, towing the damaged vehicles, and restoring traffic (order 1 hour). Sentence (9.14) can even mean the year of the accident as in "At the time of the accident John was a Junior in College." The disambiguation of the above expression requires detailed knowledge of the current context. Since the parser does not have access to this information we represent the meaning of the above phrase in an "uninterpreted" form:

```
(9.15) (TIME TIME-TYPE (NIL) SUB-REF ($VEHICLE-ACCIDENT))
```

The filler of the time-type slot must be determined from the context.

The super-referent type of time specification specifies time in terms of the time structure of the referent event. For example:

(9.16) THE TENTH WEEK OF MARY'S PREGNANCY.

In some sense Mary's pregnancy introduces its own calendar. We represent the meaning of this expression as follows:

```
(9.17) (TIME TIME-TYPE (WEEK)
             WEEK (10 SUPER-REF ($PREGNANCY
                                   ACTOR (*MARY*))))
```

The SUPER-REF link points to the event which gives meaning to the number preceding the link. In the absence of this link, default interpretations shown in Figure 9.1 are used: a month for days, a week for weekdays, a year for months, a day for dayparts, etc. When the super-referent time specification corresponds to one of the default ones, the SUPER-REF link is not used. For example:

(9.18) THE MORNING OF FEBRUARY 22, 1979.

will be represented as:

```
(9.19) (TIME TIME-TYPE  (DAYPART)
             DAYPART    (MORNING)
             DAY        (21)
             MONTH      (2)
             YEAR       (1979))
```

Here is a more complex example which combines both types of super-referent time specification:

(9.20) THE SECOND HOUR OF THE THIRD DAY OF JOHN'S VACATION.

Since a day is the default super-referent for hours, the SUPER-REF link is not used with this time-unit. We represent the meaning of the above phrase as follows:

```
(9.21) (TIME TIME-TYPE (HOUR)
             HOUR      (2)
             DAY       (3 SUPER-REF ($VACATION
                                     ACTOR (*JOHN*))))
```

The above construction still does not fully specify the meaning of (9.20) which depends on the context of the particular John's vacation. Suppose John started his vacation Monday afternoon. Is Wednesday the third or the second day of John's vacation? The answer to this question depends on how John counts time.

As we said before, the filler of a time-unit slot can be another time conceptualization. For example:

(9.22) MONDAY OF THE WEEK OF THE CAR ACCIDENT.

The car accident in the above expression did not necessarily occur on Monday. The accident is a sub-referent event for the week and not for Monday. We can represent this as follows:

```
(9.23) (TIME TIME-TYPE (DAY)
             WEEKDAY (2)
             WEEK    (TIME TIME-TYPE (WEEK)
                           SUB-REF   ($VEHICLE-ACCIDENT)))
```

Let us now examine the pre- and post-referent time specifications. In order to specify a point in time relative to some other point, we need two predicates: BEFORE and AFTER with two roles: ORIGIN and SPAN. ORIGIN gives us a reference point in time and SPAN is a distance in time. For example, the phrase "three days after X" can be represented as:

(9.24) (AFTER ORIGIN X SPAN Y)

where the time span can be represented in a way similar to date representations. We can represent "three days" as:

(9.25) (TIME-SPAN TIME-TYPE (DAYS) DAYS (3))

9.2. Parsing BY+TIME expressions.

How do we parse BY+TIME expressions into the above representations? When the parser encounters an unpredicted preposition "by" it goes to the EXPERT line of requests where the prepositional phrase expert attempts to collect the following noun group. The concept that underlies this group then triggers some of the requests stored under the preposition "by". One of these requests is the BY+TIME request which expects either the TIME concept or a number which it will try to interpret as a time specification. Let us examine the types of time expressions the noun group processor collects. We consider mostly the kinds of time specifying expressions that are normally used in newspaper stories.

There are five classes of constructions that the prepositional phrase expert can collect which will trigger one of the time processing experts:

ABS-TIME - absolute time specifications as in "5am Monday, February 26, 1979", or "Spring 1979". Absolute time specifications are handled by the ABS-TIME expert.

DEF-UNSPEC-TIME - definite but unspecified time as in "the day", "the week", etc. When the parser encounters this type of time BY+TIME construction it should normally expect an "of" prepositional phrase introducing the sub-referent event, as in "the day of John's exam". This type of constructions is handled by the SUB-REF expert.

DEF-SPEC-TIME - definite and specific time as in "the first day" or "the last week". When the parser encounters this type of BY+TIME construction it should normally expect an "of" prepositional phrase introducing a super-referent event, as in "the first day of John's vacation". This is done by the SUPER-REF expert.

INDEF-TIME - indefinite time specifications such as "a day", "a week", "three years" used in BY+TIME expression usually indicate either pre- or post-referent types of time specifications. For example: "by a week from today", or "have this done by three days before I leave!" This type of construction is handled by the BY+RELTIME expert which looks for a "from", "before", or "after" prepositional phrase introducing the referent event.

NUMBER - A noun group consisting of a single number in a BY+TIME construction usually specifies either the hour of some event (if the number is less or equal than 12) or the year. For example: "John was back by five", or "The war was over by 1945". This type of construction is handled by the

BY+NUMBER expert.

The following example illustrates the parsing of BY+TIME expressions:

(9.26) MARY EXPECTED JOHN BY 5PM MONDAY, FEBRUARY 26, 1979.

When the preposition "by" is reached, the prepositional phrase expert attempts to collect the group following the preposition. The expression 5PM builds the following structure:

```
(9.27) (TIME TIME-TYPE (HOUR) HOUR    (17)
                              WEEKDAY (NIL)
                              DAY     (NIL))
```

Each of the unfilled slots in the above conceptualization has a FORWARD directed expectation for the fillers these slots. Note that there are only two expectations active at this time: for the day of the month and for the weekday. The parser does not expect to hear specifications of the month, season or year without specifying the intermediate time-units. We do not say "5pm Spring 1979".

Next the parser reads MONDAY, which builds a DAY-type TIME conceptualization:

```
(9.28) (TIME TIME-TYPE (DAY) WEEKDAY (2)
                             DAY     (NIL))
```

We still do not have an expectation for the month because we cannot specify the month without specifying which monday is meant. The above conceptualization is picked up by the WEEKDAY expectation which produces the following structure:

```
(9.29) (TIME TIME-TYPE (HOUR) HOUR    (17)
                              WEEKDAY (2)
                              DAY     (NIL))
```

The next word is February. Days of the month in English are usually specified by the name of the month followed by the number of the day. Numbers can play a multitude of roles in a sentence but preceding the number with the name of a month sets up a specific expectation which uniquely determines a time interpretation. Thus, every name of the months in English can specify either a MONTH-type or a DAY-type time conceptualization. Names of the months have immediate WORD line expectations looking for a number following the name of the month. Thus, "FEBRUARY 26" builds the following structure:

```
(9.30) (TIME TIME-TYPE (DAY) DAY     (26)
                             MONTH   (2)
                             SEASON  (WINTER)
                             YEAR    (NIL))
```

The above conceptualization satisfies the expectation of the DAY slot in (9.29). This expectation merges the two producing:

```
(9.31) (TIME TIME-TYPE (DAY) HOUR     (17)
                             WEEKDAY  (2)
                             DAY      (26)
                             MONTH    (2)
                             SEASON   (WINTER)
                             YEAR     (NIL))
```

The only active expectation left at this point is the one attached to the YEAR slot. It picks up 1979 and fills the YEAR slot.

The following diagrams summarize the work of the BY+TIME and BY+NUMBER experts:

```
                    BY+TIME expression
                             |
                             V
                       |----------|
                       | BY+TIME  |
                       | expert   |
                       |----------|
            ABS-TIME    |  |  |  |    INDEF-TIME
  |-------------------|  |  |  |-----------------|
  |                      |  |                    |
  V          DEF-UNSPEC  |  | DEF-SPEC           V
|----------|    TIME     |  | TIME     |------------|
| ABS-TIME |  |----------|  |          | BY+RELTIME |
| expert   |  |             |          | expert     |
|----------|  |             |          |------------|
              V             V            |          |
       |---------|  |-----------|        |          |
       | SUB-REF |  | SUPER-REF |        V          |
       | expert  |  | expert    |   PRE-REFERENT    |
       |---------|  |-----------|   construction    |
                                                    |
                                                    V
                                              POST-REFERENT
                                              construction
```

Figure 9.4: BY+TIME expert

```
                    BY+NUMBER expression
                             |
                             V
                       |-----------|
                       | BY+NUMBER |
                       | expert    |
                       |-----------|
      number <,= 12          |     |   1000 < number < 2100
           |-------------|     |-------------|
           |                                 |
           V                                 V
    |-------------|                  |-------------|
    | HOUR expert |                  | YEAR expert |
    |-------------|                  |-------------|
```

Figure 9.5: BY+NUMBER expert

CHAPTER 10

SUMMARY AND CONCLUSIONS

The main objective of this thesis has been to demonstrate that parsing is a process of knowledge application. We pointed out that there are four general types of knowledge used in parsing: lexical knowledge, world knowledge, linguistic knowledge, and contextual knowledge. We described a parsing mechanism capable of using this knowledge if it is properly represented in Conceptual Dependency. The mechanism is a production-like system which uses three different modes of production organization: structure-driven, position-driven, and situation driven. The first two modes were designed to handle specific information about individual words and their meaning. Both these modes are expectation-driven. In the structure-driven mode expectations are attached to the unfilled slots of the partially built conceptual structures. In the position-driven mode expectations are attached to the concepts themselves and work on the adjacent concepts only. The first mode is used primarily in the clause-level processing, the second in the processing of noun groups. When one of these two modes fails to account for the new input, the parser goes to the third, situation-driven mode. In this mode the parser applies its knowledge about general parsing situations such as post-nominal modification, prepositional phrase attachment, etc. The assumption behind this schema is that either the definitions of individual words drive the process of parsing by making specific predictions about the concepts underlying the input, or, when this is impossible, the situation falls into one of a few standard classes which can be handled by the appropriate experts in a top-down fashion.

The parsing mechanism described in the first part of this thesis works well for sentences which satisfy the following conditions:

1) The meaning of the sentence can be represented in Conceptual Dependency.

2) There is full and accessible representation for all knowledge structures that need to be consulted during the parsing process.

In the second part of the thesis we showed that the above conditions are not always met. In the last three chapters we used the preposition "by" as a guide in our discussion of various topics in parsing. In Chapter 6 we examined passive constructions and described how they are handled in our parsing framework. In Chapter 7 we examined instrumental constructions and pointed out that their handling requires detailed knowledge of the structure of episodes in memory. In Chapter 8 we examined how "by" prepositional phrases can be used to give some additional information about the creation of physical and mental objects which usually involves their creators' ROLE-THEMES. The processing of these expressions requires detailed knowledge of ROLE-THEMES and mechanisms for their representation. We also discussed how "by" prepositional phrases contribute to the description of physical setting of the current episode. We examined some of the problems in the representations of such elements of physical settings as locations, and time. We showed that the incorporation of these elements of physical setting into the representation of texts requires detailed knowledge of the internal structure of settings. We also showed ways to represent it within the overall structure of episodes.

In many cases we were unable to give any precise algorithms and solutions. In every such case, the stumbling block was our poor understanding of the knowledge structures and memory processes involved and insufficient representational support. We cannot parse more than we can represent. More work needs to be done in such areas as representation of episodes, physical settings, people's goals and plans for their achievement, role-themes, and physical and mental objects and their functions before we can specify how to parse natural language expressions referring to these structures.

Our work on parsing follows the work on memory and representation at the Yale AI Project in a breadth-first fashion. It makes little sense in our view to try to handle narrow classes of natural language expressions outside of the general context in which they are used. A well known example is the problem of pronominal reference specification. Very complex solutions were proposed to deal with this problem on the basis of syntactic considerations (Ross 1967). However, when this problem is attacked within a system capable of contextual understanding, a large part of the problem has a relatively simple solution (Charniak 1972, Cullingford 1978). Another example of this type of problem is the problem of time keeping. In Chapter 9 we pointed out that the current episodic context often creates its own time structure. It is very difficult to parse time specifications in these cases using only absolute time scales. On the other hand, the problem might have a relatively simple solution if handled within a larger episodic memory framework.

Parsing requires application of very large amounts of knowledge. Thus, a very serious problem of knowledge management arises. The bulk of our parser's knowledge is encoded in the form of requests which embody parsing rules. There are two sources of requests in our parser: the dictionaries, which contain the majority of requests, and the experts, which organize relatively few

general purpose requests. We use the following request management mechanisms:

1. Dictionary requests can be activated by the input words.

2. Expert requests can be activated by the experts.

3. Requests know about other requests and can explicitly activate and deactivate each other.

4. Requests attached to the gaps in the conceptual structures are deactivated when the gaps are filled.

5. In noun group processing requests are activated and deactivated depending on the position of the processing node.

6. Requests are tested hierarchically depending on the request line on which they are kept.

7. Within the same line, the requests are tested in the reverse chronological order, i.e., the most recent ones are tried first.

The above mechanisms cut the number of active requests that have to be considered by the parser at any given time to a very manageable number. (In practice this number rarely exceeds half a dozen.)

From the point of view of knowledge engineering our system is not very "clean". Very few constraints are placed on the form of the requests, the requests can have all kinds of side effects, they know about each other and can interact in a variety of ways. The effects of request triggering are not easily traced. However, we still know too little about the problem to constrain the form of our rules in any meaningful way. Constraining the form of rules before their content has been sufficiently investigated might lead to unnecessary complications and waste of effort later. Our methodology is gradual abstraction of a "cleaner" system of rules as we learn more about the problem.

The main difference between our approach and most previous works in this area is our rejection of the syntax/semantics modularization. The best known parsers (Woods and Kaplan 1971, Winograd 1972) separate language analysis into two stages: syntactic and semantic. At the first stage they obtain a complete syntactic description of the sentence which is interpreted semantically at the second stage. Inefficiency and non-extendability of these programs indicate that this is a wrong kind of modularization. Moreover, complete syntactic description of a sentence is often impossible without understanding the whole sentence. Programs which rely on such modularization either cannot handle such sentences or produce several parses all but a few of which are later rejected by the semantics module. As the programs mentioned above demonstrate, complete syntactic analysis of a sentence is a difficult and very complex process (when at all possible). Not all of the syntactic information obtained in this

process is used at the later stages of understanding, but purely syntactic parsers have no way of knowing which information will be relevant and which will not. In addition to this, understanders which rely on syntactic parsers cannot deal with grammatically incorrect input that would not cause any difficulties for people. Our program demonstrates that integration of syntax and semantics is a viable alternative to the above modularization. It uses semantics to drive parsing and uses syntax only when needed and as much as needed.

The most recent of the syntactic parsers is called Parsifal and was developed by Marcus (1977). One of the claims made about Parsifal is that it can provide a simple and efficient front end for any language understanding system by producing annotated surface structures of English sentences. However, in practice, Parsifal is simple or efficient only in comparison to other syntactic parsers. In fact, its complexity (in terms of vocabulary, the number of rules, and types of constructions it can handle) is comparable to that of ELI-2. However, ELI-2 extracts meaning out of sentences, while Parsifal does only the first and easy (according to Marcus) part of the process. In addition to many of the shortcomings of the syntactic parsers mentioned above, it remains to be seen whether annotated surface structures produced by Parsifal can be usefully employed by any extensive natural language system at all. From this point of view Parsifal represents a step backward as far as natural language <u>understanding</u> is concerned.

It is difficult to talk about differences between the parser described in this thesis (ELI-2) and ELI (Riesbeck and Schank 1976) since most of ELI is an integral part of ELI-2. It would be more accurate to talk about significant extensions of the basic framework in which ELI was written. Following are three areas of such extensions which we consider most important:

1. The hierarchical organization of requests which corresponds to the natural hierarchy of tasks in parsing.

2. The position-driven mode of parsing control which gives the parser the ability to handle complex noun groups.

3. The situation-driven mode of parsing control which enables the parser to handle such a wide variety of natural language constructions as special modifiers, appositives, prepositional and participial phrases, and relative subclauses.

The description of the parser, however, was not the only objective of this thesis. Equally important was the investigation of the limitations of our current approach to parsing. If we look at the ways the parsing mechanisms described in this thesis utilize the four kinds of knowledge used in parsing: lexical knowledge, static world knowledge, linguistic knowledge, and contextual knowledge, we will see that the first three types are used much more extensively than the fourth type. The reason for this is that the parser described in this thesis is essentially dictionary-driven.

The first three types of knowledge are static and, therefore, can be stored in the parser's dictionaries. This knowledge is transformed into expectations which process the input. The ELI-2 parser represents an integrated process model of application of lexical, linguistic, and a great deal of static world knowledge. Contextual knowledge, on the other hand, is inherently dynamic and, therefore, cannot be kept in the dictionaries. Ideally, in the present parsing framework interaction between the parser and the memory which keeps contextual knowledge is supposed to occur as follows: First, the parser reads the input sentence and produces a conceptual dependency representation of its "immediate meaning". For example, the immediate meaning of:

(10.1) THE WAITER CAME TO THE TABLE.

is

```
(10.2) (PTRANS ACTOR  (*WAITER*)
               OBJECT (*WAITER*)
               TO     (PROX OBJECT (*TABLE*)))
```

This conceptualization is then handed over to memory which incorporates it in the current context. In our example, memory figures out whether the waiter came to take the order, or brought the food, etc. In this framework the parser is a separate and relatively independent module which can ask memory questions if necessary, but otherwise is driven by its dictionaries. Figure 10.1 shows this framework schematically:

```
                |----------------------------------|
                V                                  V
             |--------|                        |--------|
  input ---->| PARSER |--> immediate meaning -->| MEMORY |
             |--------|                        |--------|
                ^                                  |
                |                                  |
          |-------------|                          |
          | DICTIONARIES |<------------------------|
          |-------------|
```

Figure 10.1: Current Parsing Framework

With the increase in complexity of parsing tasks it is becoming more and more difficult to maintain this modularization. The notion of immediate meaning is becoming very fuzzy. All this indicates that separation of the parser and memory is the wrong kind of modularization just as the separation of syntax and semantics was.

The only type of interaction between the context and the parsing process in the current implementation is the use of specialized dictionaries to give preferred meanings to the words specific to the current context. As the examples in the second part

of the thesis show, context plays a much greater role in understanding than can be captured by a purely dictionary-driven parser. The parser has to be more context-driven. It should read a few words in a sentence, jump to conclusions about its possible meaning on the basis of the current context, translate these conclusions into active expectations and continue analysis in a top-down fashion. This is the area in which we intend to concentrate in our future research. Thus, we view the work described in this thesis as the first step (from the parsing side) toward an integrated model of natural language understanding.

BIBLIOGRAPHY

1. Birnbaum, L. A. and Selfridge M. G. (1979). On Conceptual Analysis of English. Computer Science Research Report (forthcoming). Yale University, New Haven, CT.

2. Bruce, B. (1972). A Model for Temporal References and its Application in a Question Answering Program. Artificial Intelligence, Vol. 3, No. 1.

3. Carbonell, J. G., Cullingford, R. E., and Gershman A. V. (1978). Knowledge-Based Machine Translation. Computer Science Research Report No. 146. Yale University, New Haven, CT.

4. Carbonell, J. G. (1979). Subjective Understanding: Computer Models of Belief Systems. Ph.D. Dissertation. Computer Science Research Report No. 150. Yale University, New Haven, CT.

5. Charniak, E. (1972). Towards a Model of Children's Story Comprehension. (PhD. thesis) AITR-266, M.I.T., Cambridge, Mass.

6. Charniak, E. (1978). A Framed Painting -- Representation of a Commonsense Knowledge Fragment. Cognitive Science, Vol. 2, No. 1.

7. Chomsky, N. (1973). Conditions on Transformations. In S. Anderson and P. Kiparsky, eds., A Festschrift for Morris Halle, Holt, Rinehart and Winston, N. Y.

8. Cullingford, R. E. (1978). Script Application: Computer Understanding of Newspaper Stories. Ph.D. Dissertation. Computer Science Research Report No. 116. Yale University, New Haven, CT.

9. DeJong, G. F. (1977). Skimming Newspaper Stories by Computer. Computer Science Research Report No. 104. Yale University, New Haven, CT.

10. Gershman, A. (1977). Analyzing English Noun Groups for Their Conceptual Content. Computer Science Research Report No. 110. Yale University, New Haven, CT.

11. Granger, R.H. (1977). FOUL-UP. Proceedings of the Fifth International Joint Conference on Artificial Intelligence. Cambridge, Massachusetts.

12. Kahn, K. and Gorry, G. A. (1977). Mechanizing Temporal Knowledge. Artificial Intelligence, Vol. 9, No. 1.

13. Kolodner, J. L. (1978). Memory Organization for Natural Language Data-Base Inquiry. Computer Science Research Report No. 142. Yale University, New Haven, CT.

14. Lehnert, W. (1977). The Process of Question Answering. (PhD. thesis). Department of Computer Science Report No. 88. Yale University, New Haven, CT.

15. Marcus, M. (1977). Theory of Syntactic Recognition for Natural Language. Ph.D. Dissertation. MIT, Cambridge, Massachusetts.

16. Minsky, M. (1975). A Framework for Representing Knowledge. _The Psychology of Computer Vision_. P. H. Winston (ed.), McGraw-Hill, New York.

17. Newell, A. (1973). Production systems: Models of Control Structures. In Chase, W.C. ed. _Visual Information Processing_. Academic Press, New York.

18. Riesbeck, C. (1974). Conceptual Understanding: Analysis of Sentences in Context. Ph.D. Dissertation. Computer Science Research Report. Stanford University, Stanford, CA.

19. Riesbeck, C. (1975). Conceptual Analysis. _Conceptual Information Processing_. R. Schank (ed.), North-Holland, Amsterdam.

20. Riesbeck, C. and Schank, R. (1976). Comprehension by Computer: Expectation-Based Analysis of Sentences in Context. Research Report #78. Department of Computer Science, Yale University, New Haven, Ct.

21. Riesbeck, C. and Charniak, E. (1978). Micro-SAM and Micro ELI: Exercises in Popular Cognitive Mechanics. Research Report #139. Department of Computer Science, Yale University, New Haven, Ct.

22. Ross, J. R. (1967). On the Cyclic Nature of English Pronominalization. In _To Honor Roman Jakobson_. Mouton Co. The Hague, Netherland.

23. Russell, S. G. (1975). Computer Understanding of Conceptually Complex Phrases. Ph.D. Dissertation. Computer Science Research Report. Stanford University, Stanford, CA.

24. Schank, R. C. (1974). Understanding Paragraphs. Technical Report #6. Instituto per gli Studi Semantici e Cognitivi, Castagnola, Switzerland.

25. Schank, R. C. (1975). _Conceptual Information Processing_. (ed.) North-Holland, Amsterdam.

26. Schank, R. C. and Abelson, R. P. (1977). _Scripts, Plans, Goals and Understanding_. Lawrence Erlbaum Assoc., Hillsdale, N.J.

27. Schank, R. C. and Colby, K. M. (1973). Computer Models of Thought and Language. W. H. Freeman and Co., San Francisco.

28. Wilensky, R. (1978). Understanding Goal-Based Stories. Ph.D. Dissertation. Computer Science Research Report No. 140. Yale University, New Haven, CT.

29. Wilks, Y. (1973). An Artificial Intelligence Approach to Machine Translation. In R. Schank and K. Colby (eds.), Computer Models of Thought and Language, W. H. Freeman and Co., San Francisco.

30. Wilks, Y. (1976). A Preferential, Pattern-Seeking, Semantics for Natural Language Inference. Artificial Intelligence, Vol. 6.

31. Winograd, T. (1972). Understanding Natural Language. Academic Press, New York.

32. Woods, W. A. (1970). Transition Network Grammars for Natural Language Analysis. Comm. ACM. Vol. 13, no. 10, 591-606.

33. Woods, W.A. and Kaplan, R.M. (1971). The Lunar Sciences Natural Language Information System. BBN Report No. 2265. Bolt Beranek and Newman Inc. Cambridge, Massachusetts.

34. Woods, W. A., et. al. (1972). The Lunar Sciences Natural Language Information System: Final Report, Report No. 2378. Bolt, Beranek and Newman, Cambridge, Massachusetts.